Chapter 2
Formulation Techniques Involving Transformations of Variables

2.1 Operations Research: The Science of Better

Operations Research (OR) is the branch of science dealing with tools or techniques for decision making to optimize the performance of systems, that is, to make those systems better. Measures of performance, of which there may be several, are numerical criteria that gauge the quality of some aspect of system’s performance, for example, annual profit or market share of a company, etc. They are of two types: (1) profit measures: (for these, the higher the value the better), (2) cost measures: (for these the lower the value the better).

OR deals with techniques for designing ways to operate the system to maximize profit measures or minimize cost measures as desired. Hence OR is the science to make systems better.

Linear Programming (LP) is an important branch of OR dealing with decision problems modeled as those of optimizing a linear function of decision variables subject to linear constraints that may include equality constraints, inequality constraints, and bounds in decision variables. In an LP, all decision variables are required to be continuous variables that can assume all possible values within their bounds subject to the constraints. LPs are special instances of mathematical programming. Besides LP, the subject mathematical programming includes network, integer, combinatorial, discrete, quadratic, and nonlinear programming.

The focus of this book is to study important aspects of LP and QP (quadratic programming) and their intelligent applications for decision making.

We refer the reader to Chap. 3 in the Junior-level book (Murty (2005b) of Chap. 1; this book can be downloaded from the website mentioned there), where decision-making problems that can be modeled directly as LPs are discussed with many illustrative examples. In this chapter we extend the range of applications of LP to include decision-making problems involving the optimization of a piecewise linear objective function subject to linear constraints. When the objective function satisfies certain properties, these problems can be transformed into LPs in terms of additional variables.
2.2 Differentiable Convex and Concave Functions

The concepts of convexity of functions, and of sets, are fundamental pillars in optimization theory. We already know that a subset \( K \subseteq \mathbb{R}^n \) is said to be a **convex set** if for every pair of points \( x, y \in K \), every convex combination of \( x, y \) (i.e., point of the form \( \alpha x + (1 - \alpha)y \) for any \( 0 \leq \alpha \leq 1 \)) is also in \( K \).

A real-valued function \( f(x) \) of decision variables \( x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n \) is said to be a **linear function** if it satisfies the following two properties that together are known as the **linearity assumptions**:  

- **Proportionality:** \( f(\alpha x) = \alpha f(x) \) for all \( x \in \mathbb{R}^n, \alpha \in \mathbb{R}^1 \)
- **Additivity:** \( f(x + y) = f(x) + f(y) \) for all \( x, y \in \mathbb{R}^n \)

An equivalent definition is: The real-valued function \( f(x) \) defined over \( x \in \mathbb{R}^n \) is a linear function, iff there exists a row vector of constants \( c = (c_1, \ldots, c_n) \) such that \( f(x) = c_1 x_1 + \ldots + c_n x_n = cx \) for all \( x \in \mathbb{R}^n \). In fact, for each \( j = 1 \) to \( n \),  
\[ c_j = f(I_j), \]  
where \( I_j \) is the \( j \)th column vector of the unit matrix \( I \) of order \( n \).

A real-valued function \( \theta(x) \) of decision variables \( x \in \mathbb{R}^n \) is said to be an **affine function** if there exists a constant \( c_0 \) such that \( \theta(x) - c_0 \) is a linear function as defined earlier. Actually this constant \( c_0 = \theta(0) \). Thus equivalently, \( \theta(x) \) is an affine function iff there exist constants \( c_0, c_1, \ldots, c_n \) such that \( \theta(x) = c_0 + c_1 x_1 + \ldots + c_n x_n \).

The concept of convexity of a function is defined by **Jensen’s inequality** stated below; it is related to the concept of convexity of a set, but we will not discuss this relationship in this book as it is not important for the things we discuss here. A function is said to be concave if its negative is convex, but there is no corresponding concept called “concavity” for sets.

Linear and affine functions are both convex and concave; but convex and concave functions may be nonlinear. In this section, we study important properties of differentiable convex, concave functions, which may be nonlinear. A requirement is that the set on which a convex or concave function is defined must be a convex set. We will study convex, concave functions defined over \( \mathbb{R}^n \) (or over a convex subset of it) for \( n \geq 1 \) in this section.

### 2.2.1 Convex and Concave Functions

A real-valued function \( g(y) \) defined over some convex subset \( \Gamma \subseteq \mathbb{R}^n \) (\( \Gamma \) may be \( \mathbb{R}^n \) itself) is said to be a **convex function** if  
\[ g(\alpha y^1 + (1 - \alpha)y^2) \leq \alpha g(y^1) + (1 - \alpha)g(y^2) \]

for all \( y^1, y^2 \in \Gamma \), and \( 0 \leq \alpha \leq 1 \). This inequality defining a convex function is called **Jensen’s inequality** after the Danish mathematician who introduced it.
To interpret Jensen’s inequality geometrically, introduce an \((n + 1)\)th axis for plotting the function value. So points in this space \(R^{n+1}\) are \((y, y_{n+1})^T\), where on the \(y_{n+1}\)th axis we plot the function value \(g(y)\) to get a geometric representation of the function.

The set of all points \(\{(y, g(y))^T : y \in \Gamma\}\) in this space \(R^{n+1}\) is a surface, which is the surface or graph of the function \(g(y)\).

The line segment \(\{(\alpha y^1 + (1 - \alpha)y^2, \alpha g(y^1) + (1 - \alpha)g(y^2))^T : 0 \leq \alpha \leq 1\}\) joining the two points \((y^1, g(y^1))^T, (y^2, g(y^2))^T\) on the graph of the function is called the chord of the function between the points \(y^1, y^2\) or on the one-dimensional line interval joining \(y^1\) and \(y^2\). If we plot the function curve and the chord on the line segment \(\{\alpha y^1 + (1 - \alpha)y^2 : 0 \leq \alpha \leq 1\}\), then Jensen’s inequality requires that the function curve lie beneath the chord. See Fig. 2.1 where the function curve and a chord are shown for a function \(\theta(\lambda)\) of one variable \(\lambda\).

The real-valued function \(h(y)\) defined on a convex subset \(\Gamma \subset R^n\) is said to be a concave function if \(-h(y)\) is a convex function, that is, if

\[
h(\alpha y^1 + (1 - \alpha)y^2) \geq \alpha h(y^1) + (1 - \alpha)h(y^2)
\]

for all \(y^1, y^2 \in \Gamma\) and \(0 \leq \alpha \leq 1\); see Fig. 2.2. For a concave function \(h(y)\), the function curve always lies above every chord.

**Fig. 2.1** Graph of a convex function \(\theta(\lambda)\) defined on \(R^1\) and its chord between two points \(\lambda_1\) and \(\lambda_2\)

**Fig. 2.2** Graph of a concave function \(\theta(\lambda)\) defined on \(R^1\) and its chord between two points \(\lambda_1\) and \(\lambda_2\)
All linear and affine functions (i.e., functions of the form $cx + c_0$, where $c \in R^n$, $c_0 \in R^1$ are given, and $x \in R^n$ is the vector of variables) are both convex and concave.

Other examples of convex functions are $\lambda^2$, $e^\lambda$ over $\lambda \in R^1$, where $r$ is a positive integer; $- \log(\lambda)$ over $\{\lambda > 0 : \lambda \in R^1\}$; and the quadratic function $x^T Dx + cx + c_0$ over $x \in R^n$, where $D$ is a positive semidefinite (PSD) matrix of order $n$ (a square matrix $D$ of order $n \times n$ is said to be a PSD (positive semidefinite) matrix iff $x^T Dx \geq 0$ for all $x \in R^n$. See Kaplan (1999); Murty (1988, 1995), or Sect. 9.1 for discussion of positive semidefiniteness of a square matrix, and the proof that this quadratic function is convex over the whole space $R^n$ iff $D$ is PSD).

We now derive some important properties of differentiable convex, concave functions. For this discussion, the functions may be nonlinear.

**Theorem 2.1. Gradient support inequality for convex functions:** Let $g(y)$ be a real-valued differentiable function defined on $R^n$. Then $g(y)$ is a convex function iff

$$g(y) \geq g(\bar{y}) + \nabla g(\bar{y})(y - \bar{y})$$

for all $y$, $\bar{y} \in R^n$, where $\nabla g(\bar{y}) = \left(\frac{\partial g(\bar{y})}{\partial y_1}, \ldots, \frac{\partial g(\bar{y})}{\partial y_n}\right)$ is the row vector of partial derivatives of $g(y)$ at $\bar{y}$.

**Proof.** Assume that $g(y)$ is convex. Let $0 < \alpha < 1$. Then $(1 - \alpha)\bar{y} + \alpha y = \bar{y} + \alpha(y - \bar{y})$. So, from Jensen’s inequality $g(\bar{y} + \alpha(y - \bar{y})) \leq (1 - \alpha)g(\bar{y}) + \alpha g(y)$. So

$$g(y) - g(\bar{y}) \geq \frac{g(\bar{y} + \alpha(y - \bar{y})) - g(\bar{y})}{\alpha}.$$

Taking the limit as $\alpha \to 0$, by the definition of differentiability, the RHS in the above inequality tends to $\nabla g(\bar{y})(y - \bar{y})$. So we have $g(y) - g(\bar{y}) \geq \nabla g(\bar{y})(y - \bar{y})$.

Now suppose the inequality in the statement of the theorem holds for all points $\bar{y}$, $y \in R^n$. Let $y^1$, $y^2$ be any two points in $R^n$ and $0 < \alpha < 1$. Taking $y = y^1$, $\bar{y} = (1 - \alpha)y^1 + \alpha y^2$, we get the first inequality given below; and taking $y = y^2$, $\bar{y} = (1 - \alpha)y^1 + \alpha y^2$, we get the second inequality given below.

$$g(y^1) - g(((1 - \alpha)y^1 + \alpha y^2) \geq \alpha(\nabla g(((1 - \alpha)y^1 + \alpha y^2)(y^1 - y^2),$$

$$g(y^2) - g(((1 - \alpha)y^1 + \alpha y^2) \geq -(1 - \alpha)(\nabla g(((1 - \alpha)y^1 + \alpha y^2)(y^1 - y^2).$$

Multiplying the first inequality above by $(1 - \alpha)$ and the second by $\alpha$ and adding, we get $(1 - \alpha)g(y^1) + \alpha g(y^2) - g((1 - \alpha)y^1 + \alpha y^2) \geq 0$, which is Jensen’s inequality. As this holds for all $y^1, y^2 \in R^n$ and $0 < \alpha < 1$, $g(y)$ is convex by definition. □
2.2 Differentiable Convex and Concave Functions

At any given point \( \bar{y} \), the function \( L(y) = g(\bar{y}) + \nabla g(\bar{y})(y - \bar{y}) \) is an affine function of \( y \), which is known as the linearization of the differentiable function \( g(y) \) at the point \( \bar{y} \). Theorem 2.1 shows that for a differentiable convex function \( g(y) \), its linearization \( L(y) \) at any point \( \bar{y} \) is an underestimate for \( g(y) \) at every point \( y \); see Fig. 2.3.

The corresponding result for concave functions obtained by applying the result in Theorem 2.1 to the negative of the function is given in Theorem 2.2.

**Theorem 2.2. Gradient support inequality for concave functions:** Let \( h(y) \) be a real-valued differentiable function defined on \( \mathbb{R}^n \). Then \( h(y) \) is a concave function iff

\[
h(y) \leq h(\bar{y}) + \nabla h(\bar{y})(y - \bar{y})
\]

for all \( y, \bar{y} \in \mathbb{R}^n \), where \( \nabla h(\bar{y}) = \left( \frac{\partial h(\bar{y})}{\partial y_1}, \ldots, \frac{\partial h(\bar{y})}{\partial y_n} \right) \) is the row vector of partial derivatives of \( h(y) \) at \( \bar{y} \). That is, the linearization of a concave function at any given point \( \bar{y} \) is an overestimate of the function at every point; see Fig. 2.4.

**Theorem 2.3.** Let \( \theta(y) \) be a real-valued differentiable function defined on \( \mathbb{R}^n \). Then \( \theta(y) \) is a convex [concave] function iff for all \( y^1, y^2 \in \mathbb{R}^n \)

\[
\{\nabla \theta(y^2) - \nabla \theta(y^1)\}(y^2 - y^1) \geq 0 \quad [\leq 0].
\]

**Proof.** We will give the proof for the convex case, and the concave case is proved similarly.

Suppose \( \theta(y) \) is convex, and let \( y^1, y^2 \in \mathbb{R}^n \). From Theorem 2.1 we have

\[
\theta(y^2) - \theta(y^1) - \nabla \theta(y^1)(y^2 - y^1) \geq 0,
\]

\[
\theta(y^1) - \theta(y^2) - \nabla \theta(y^2)(y^1 - y^2) \geq 0.
\]
Adding these two inequalities, we get \( \{ \nabla \theta(y^2) - \nabla \theta(y^1) \}(y^2 - y^1) \geq 0 \).

Now suppose that \( \theta(y) \) satisfies the property stated in the theorem; and let \( y^1, y^2 \in \mathbb{R}^n \). As \( \theta(y) \) is differentiable, by the mean-value theorem of calculus, we know that there exists an \( 0 < \alpha < 1 \) such that \( \theta(y^2) - \theta(y^1) = \nabla \theta(y^1 + \alpha(y^2 - y^1))(y^2 - y^1) \). As \( \theta(y) \) satisfies the statement in the theorem, we have

\[
\{ \nabla \theta(y^1 + \alpha(y^2 - y^1)) - \nabla \theta(y^1) \} \alpha(y^2 - y^1) \geq 0 \quad \text{or} \quad \nabla \theta(y^1 + \alpha(y^2 - y^1))(y^2 - y^1) \geq \nabla \theta(y^1)(y^2 - y^1).
\]

But by the choice of \( \alpha \) as discussed above, the left-hand side of the last inequality is \( = \theta(y^2) - \theta(y^1) \). Therefore, \( \theta(y^2) - \theta(y^1) \geq \nabla \theta(y^1)(y^2 - y^1) \). Since this holds for all \( y^1, y^2 \in \mathbb{R}^n \), by Theorem 2.1, \( \theta(y) \) is convex.

Applying Theorem 2.3 to a function defined over \( \mathbb{R}^1 \), we get the following result:

**Result 2.1.** Let \( \theta(\lambda) \) be a differentiable real-valued function of a single variable \( \lambda \in \mathbb{R}^1 \). \( \theta(\lambda) \) is convex [concave] iff its derivative \( \frac{d\theta}{d\lambda} \) is a monotonic increasing [decreasing] function of \( \lambda \).

Hence checking whether a given differentiable function of a single variable \( \lambda \) is convex or concave involves checking whether its derivative is a monotonic function of \( \lambda \). If the function is twice continuously differentiable, this will hold if the second derivative has the same sign for all \( \lambda \). If the second derivative is \( \geq 0 \) for all \( \lambda \), the function is convex; if it is \( \leq 0 \) for all \( \lambda \), the function is concave.

Now we will discuss the generalization of Result 2.1 to functions defined on \( \mathbb{R}^n \) for \( n \geq 2 \). A square matrix \( D \) of order \( n \) is said to be positive [negative]
2.2 Differentiable Convex and Concave Functions

semidefinite (PSD or [NSD]) if $x^T Dx \geq 0$ for all $x \in \mathbb{R}^n$. In Chap. 9 these concepts are defined and efficient algorithms for checking whether a given square matrix satisfies these properties are discussed.

**Theorem 2.4.** Let $g(y)$ be a twice continuously differentiable real-valued function defined on $\mathbb{R}^n$, and let $H(g(y)) = \left( \frac{\partial^2 g(y)}{\partial y_i \partial y_j} \right)$ denote its Hessian matrix (the $n \times n$ matrix of second partial derivatives) at $y$. Then $g(y)$ is convex iff $H(g(y))$ is a PSD (positive semi-definite) matrix for all $y$. Correspondingly, $g(y)$ is concave iff $H(g(y))$ is a NSD (negative semi-definite) matrix for all $y$.

**Proof.** We will prove the convex case. Consider a point $\tilde{y} \in \mathbb{R}^n$.

Suppose $g(y)$ is convex. Let $\alpha > 0$ and sufficiently small. By Theorem 2.1 we have for each $x \in \mathbb{R}^n$

$$(g(\tilde{y} + \alpha x) - g(\tilde{y}) - \alpha \nabla g(\tilde{y}) x)/\alpha \geq 0$$

Take limit as $\alpha \to 0^+$ (through positive values of $\alpha$). By the mean value theorem of calculus the left-hand side of the above inequality converges to $x^T H(g(\tilde{y})) x$, and hence we have $x^T H(g(\tilde{y})) x \geq 0$ for all $x \in \mathbb{R}^n$, this is the condition for the Hessian matrix $H(g(\tilde{y}))$ to be PSD.

Suppose $H(g(y))$ is PSD for all $y \in \mathbb{R}^n$. Then by Taylor’s theorem of calculus, for any $y_1, y_2 \in \mathbb{R}^n$

$$g(y_2) - g(y_1) - \nabla g(y_1)(y_2 - y_1) = (y_2 - y_1)^T H(g(y_1 + \alpha(y_2 - y_1)))(y_2 - y_1)$$

for some $0 < \alpha < 1$, which is $\geq 0$ since $H(g(y_1 + \alpha(y_2 - y_1)))$ is PSD. So the right-hand side of the above equation is $\geq 0$ for all $y_1, y_2 \in \mathbb{R}^n$; therefore $g(y)$ is convex by Theorem 2.1. \(\square\)

We know that linear and affine functions are both convex and concave. Now consider the general quadratic function $f(x) = x^T D x + c x + c_0$ in variables $x \in \mathbb{R}^n$, its Hessian matrix $H(f(x)) = (D + D^T)/2$ is a constant matrix. Hence the quadratic function $f(x)$ is convex iff the matrix $(D + D^T)/2$ is a PSD matrix by Theorem 2.4. Checking whether a given square matrix of order $n$ is PSD can be carried out very efficiently with an effort of at most $n$ Gaussian pivot steps (see Kaplan (1999); Murty (1988), or Sect. 9.2 of this book, for the algorithm to use). So whether a given quadratic function is convex or not can be checked very efficiently.

For checking whether a general twice continuously differentiable nonlinear function of $x$ outside the class of linear and quadratic functions is convex may be a hard problem, because its Hessian matrix depends on $x$, and the job requires checking that the Hessian matrix is a PSD matrix for every $x$. Fortunately, for piecewise linear (PL) functions, which we will discuss in the next section, checking whether they are convex can be carried out very efficiently even though those functions are not differentiable everywhere.
2.3 Piecewise Linear (PL) Functions

**Definition: Piecewise Linear (PL) Functions:** Considering real-valued continuous functions $f(x)$ defined over $\mathbb{R}^n$, these are nonlinear functions that may not satisfy the linearity assumptions over the whole space $\mathbb{R}^n$, but there is a partition of $\mathbb{R}^n$ into convex polyhedral regions, say $\mathbb{R}^n = K_1 \cup K_2 \cup \ldots \cup K_r$ such that $f(x)$ is an affine function within each of these regions individually, that is, for each $1 \leq t \leq r$

there exist constants $c_i^t, c^t \equiv (c_1^t, \ldots, c_n^t)$ such that $f(x) = f_t(x) = c_0^t + c^t x$ for all $x \in K_t$, and for every $S \subset \{1, \ldots, r\}$, and at every point $x \in \cap_{t \in S} K_t$, the different functions $f_t(x)$ for all $t \in S$ have the same value.

Now we give some examples of continuous PL functions defined over $\mathbb{R}^1$. Denote the variable by $\lambda$.

Each convex polyhedral subset of $\mathbb{R}^1$ is an interval; so a partition of $\mathbb{R}^1$ into convex polyhedral subsets expresses it as a union of intervals:

$\lambda_1 \leq \lambda \leq \lambda_2 = \{\lambda_1 \leq \lambda \leq \lambda_2\}, \ldots, [\lambda_{r-1}, \lambda_r], [\lambda_r, \infty]$, where $\lambda_1, \ldots, \lambda_r$ are the boundary points of the various intervals, usually called the breakpoints in this partition.

The function $\theta(\lambda)$ is a PL function if there exists a partition of $\mathbb{R}^1$ like this such that inside each interval of this partition the slope of $\theta(\lambda)$ is a constant, and its value at each breakpoint agrees with the limits of $\theta(\lambda)$ as $\lambda$ approaches this breakpoint from the left, or right; that is, it should be of the form tabulated below:

<table>
<thead>
<tr>
<th>Interval</th>
<th>Slope of $\theta(\lambda)$ in interval</th>
<th>Value of $\theta(\lambda)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda \leq \lambda_1$</td>
<td>$c_1$</td>
<td>$c_1 \lambda$</td>
</tr>
<tr>
<td>$\lambda_1 \leq \lambda \leq \lambda_2$</td>
<td>$c_2$</td>
<td>$\theta(\lambda_1) + c_2 (\lambda - \lambda_1)$</td>
</tr>
<tr>
<td>$\lambda_2 \leq \lambda \leq \lambda_3$</td>
<td>$c_3$</td>
<td>$\theta(\lambda_2) + c_3 (\lambda - \lambda_2)$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$\lambda_{r-1} \leq \lambda \leq \lambda_r$</td>
<td>$c_r$</td>
<td>$\theta(\lambda_{r-1}) + c_r (\lambda - \lambda_{r-1})$</td>
</tr>
<tr>
<td>$\lambda \geq \lambda_r$</td>
<td>$c_{r+1}$</td>
<td>$\theta(\lambda_r) + c_{r+1} (\lambda - \lambda_r)$</td>
</tr>
</tbody>
</table>

Notice that the PL function $\theta(\lambda)$ defined in the table above is continuous, and at each of the breakpoints $\tilde{\lambda} \in \{\lambda_1, \ldots, \lambda_r\}$ we verify that

$$
\lim_{\epsilon \to 0^-} \theta(\tilde{\lambda} + \epsilon) = \lim_{\epsilon \to 0^+} \theta(\tilde{\lambda} + \epsilon) = \theta(\tilde{\lambda}).
$$

Here are numerical examples of continuous PL functions:

**Example 2.1.**

<table>
<thead>
<tr>
<th>Interval</th>
<th>Slope of $\theta(\lambda)$ in interval</th>
<th>Values of $\theta(\lambda)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-\infty$ to 10</td>
<td>3</td>
<td>$3\lambda$</td>
</tr>
<tr>
<td>10–25</td>
<td>5</td>
<td>$30 + 5(\lambda - 10)$</td>
</tr>
<tr>
<td>25 to $\infty$</td>
<td>7</td>
<td>$105 + 7(\lambda - 25)$</td>
</tr>
</tbody>
</table>
Example 2.2.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Slope of (\theta(\lambda)) in interval</th>
<th>Values of (\theta(\lambda))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-\infty) to 100</td>
<td>10</td>
<td>(10\lambda)</td>
</tr>
<tr>
<td>100–300</td>
<td>5</td>
<td>(1,000 + 5(\lambda -100))</td>
</tr>
<tr>
<td>300–1,000</td>
<td>10</td>
<td>(2,000 + 10(\lambda -300))</td>
</tr>
<tr>
<td>1,000 to (\infty)</td>
<td>20</td>
<td>(9,000 + 20(\lambda -1,000))</td>
</tr>
</tbody>
</table>

Exercises

2.3.1. (1) Show that the sum of PL functions is PL. Show that a linear combination of PL functions is PL.

(2) Show that the function \(\lambda = 1/(1-\lambda)^2\) is convex on the set \(-1 \leq \lambda < 1\). Also, show that the function \(\lambda^6 - 15\lambda^2\) is convex on the set \(2 \leq \lambda \leq 3\).

2.3.2. Is the subset of \(R^2, \{x = (x_1, x_2)^T : x_1x_2 > 1\}\), a convex set? What about its complement?

2.3.3. Show that a real-valued function \(f(x)\) of decision variables \(x \in R^n\) is an affine function iff for any \(x \in R^n\) the function \(g(y) = f(x + y) - f(x)\) is a linear function of \(y\).

2.3.4. Let \(K_1 \cup K_2 \cup \ldots \cup K_r\) be a partition of \(R^n\) into convex polyhedral regions, and \(f(x)\) a real-valued continuous function defined on \(R^n\). Show that \(f(x)\) is a PL function with this partition of \(R^n\) iff it satisfies the following properties: for each \(t \in \{1, \ldots, r\}, x \in K_t\)

1. and all \(y\) such that \(x + \alpha y \in K_t\) for some \(\alpha > 0\), 
   \[f(x + \lambda y) = f(x) + \lambda ((f(x + \alpha y) - f(x))/\alpha)\] for all \(\lambda \geq 0\) such that \(x + \lambda y \in K_t\); and
2. for each \(y^1, y^2 \in R^n\) such that \(x + y^1, x + y^2\) are both in \(K_t\), if \(x + y^1 + y^2 \in K_t\) also, then 
   \[f(x + y^1 + y^2) = f(x) + (f(x + y^1) - f(x)) + (f(x + y^2) - f(x)).\]

2.3.5. Show that the function \(f(x) = (x_1^2)/(c_0 + c_1x_1 + c_2x_2)\) of \(x \in R^3\) is a convex function on the set \(\{x \in R^3 : c_0 + c_1x_1 + c_2x_2 > 0\}\).

2.3.1 Convexity of PL Functions of a Single Variable

We discuss convexity of PL functions next. As these functions are not differentiable at points where there slopes change, the arguments used in the previous section based on differentiability do not apply.

Result 2.2. Let \(\theta(\lambda)\) be a PL function of a single variable \(\lambda \in R^1\). Let \(\lambda_1, \ldots, \lambda_r\) be the various breakpoints in increasing order where its slope changes. \(\theta(\lambda)\) is
convex iff at each breakpoint $\lambda_t$ its slope to the right of $\lambda_t$ is strictly greater than its slope to the left of $\lambda_t$; that is, iff its slopes are monotonic increasing with the variable.

Proof. Suppose at a breakpoint $\lambda_t$, $c_t = \text{the slope of } \theta(\lambda)$ to the right of $\lambda_t$ is $< c_{t-1} = \text{its slope to the left of } \lambda_t$. Let $\tilde{\lambda}$ be a point close to but $< \lambda_t$, where the slope of $\theta(\lambda)$ is $c_{t-1}$, and $\tilde{\lambda}$ is a point close to but $> \lambda_t$, where its slope is $c_t$. Then the graph of $\theta(\lambda)$ in the neighborhood of $\lambda_t$ will be as shown by the solid line in Fig. 2.5. The chord of the function in the interval $\tilde{\lambda} \leq \lambda \leq \tilde{\lambda}$ shown by the dashed line segment is below the function, violating Jensen’s inequality for convex functions. So, $\theta(\lambda)$ cannot be convex.

If the slopes of the function satisfy the condition mentioned in the Result, then it can be verified that every chord lies above the function, establishing its convexity.

The corresponding result for concave functions is: a PL function of one variable is concave iff its slope to the right of every breakpoint is less than its slope to the left of that breakpoint, that is, its slopes are monotonic decreasing with the variable. These results provide a convenient way to check whether a PL function of one variable is convex, or concave, or neither. For example, the PL function in Example 2.1 has monotonically increasing slopes, so it is convex. For the one in Example 2.2, the slope is not monotone, so it is neither convex nor concave.

### 2.3.2 PL Convex and Concave Functions in Several Variables

Let $f(x)$ be a PL function of variables $x = (x_1, \ldots, x_n)^T$ defined over $R^n$. So, there exists a partition $R^n = \bigcup_{t=1}^{r} K_t$, where $K_t$ is a convex polyhedral set for all $t$, the interiors of $K_1, \ldots, K_r$ are mutually disjoint, and $f(x)$ is affine in each $K_t$; that is, we have vectors $c^t$ and constants $c^t_0$ such that

$$f(x) = c^T_0 + c^t x \quad \text{for all } x \in K_t, \ t = 1 \text{ to } r.$$  

(2.1)
Checking the convexity of \( f(x) \) on \( R^n \) is not as simple as in the one-dimensional case (when \( n = 1 \)), but the following theorem explains how it can be done.

**Theorem 2.5.** Let \( K_1 \cup \ldots \cup K_r \) be a partition of \( R^n \) into convex polyhedral regions, and \( f(x) \) the PL function defined by the above equation (2.1). Then \( f(x) \) is convex iff for each \( t = 1 \) to \( r \), and for all \( x \in K_t \)

\[
c'_0 + c'x = \text{Maximum}\{c'_0 + c^p x : \ p = 1, \ldots, r\}
\]

In effect, this says that \( f(x) \) is convex iff for each \( x \in R^n \)

\[
f(x) = \text{Maximum}\{c'_0 + c^p x : \ p = 1, \ldots, r\}
\]  

(2.2)

**Proof.** Suppose \( f(x) \) satisfies the condition (2.2) stated in the theorem. Let \( x^1, x^2 \in R^n \) and \( 0 \leq \alpha \leq 1 \). Suppose

\[
f(x^1) = \text{Maximum}\{c'_0 + c^p x^1 : \ p = 1, \ldots, r\} = c'_0 + c^1 x^1,
\]  

(2.3)

\[
f(x^2) = \text{Maximum}\{c'_0 + c^p x^2 : \ p = 1, \ldots, r\} = c'_0 + c^2 x^2,
\]  

(2.4)

and \( f(\alpha x^1 + (1 - \alpha)x^2) = \text{Maximum}\{c'_0 + c^p (\alpha x^1 + (1 - \alpha)x^2) : \ p = 1, \ldots, r\} = c'_0 + c^a (\alpha x^1 + (1 - \alpha)x^2) \) for some \( a \). Then

\[
f(\alpha x^1 + (1 - \alpha)x^2) = (a'_0 + a^a x^1) + (1 - \alpha)(a'_0 + a^a x^2),
\]

\[
\leq \alpha(c'_0 + c^1 x^1) + (1 - \alpha)(c'_0 + c^2 x^2)
\]

from (2.3), (2.4),

\[
= \alpha f(x^1) + (1 - \alpha) f(x^2).
\]

As this holds for all \( x^1, x^2 \in R^n \) and \( 0 \leq \alpha \leq 1 \), \( f(x) \) is convex by definition.

Now suppose that \( K_1 \cup \ldots \cup K_r \) is a partition of \( R^n \) into convex polyhedral regions, and \( f(x) \) the PL function defined by \( f(x) = c'_0 + c'x \) for all \( x \in K_t \), \( t = 1 \) to \( r \), is convex. Let \( \tilde{x} \) be any point in \( R^n \), suppose \( \tilde{x} \in K_{p_0} \). Let \( x^1 \in K_1 \), \( x^2 \in K_2 \) be any two points such that \( \tilde{x} \) is on the line segment \( L \) joining them, that is, \( \tilde{x} = \lambda x^1 + (1 - \lambda)x^2 \) for some \( 0 < \lambda < 1 \). For \( 0 \leq \lambda \leq 1 \) let

\[
f(\lambda x^1 + (1 - \lambda)x^2) = \theta(\lambda).
\]

The line segment \( L \) begins in \( K_{p_0} \), where \( p_0 = 1 \), and suppose it goes through \( K_{p_1}, K_{p_2}, \ldots, K_{p_{p_0}}, K_{p_{p_0} + 1}, \ldots, K_{p_s} \), where \( p_s = 2 \); this breaks up \( L \) into \( s - 1 \) intervals, each interval being the portion of \( L \) in one of the sets \( K_{p_1}, \ldots, K_{p_s} \). Let the breakpoints for these intervals be \( \lambda_1, \ldots, \lambda_s \) in increasing order.

So, in the interval \( 0 \leq \lambda \leq \lambda_1 \), \( \theta(\lambda) = c'_0 + c^1(\lambda x^1 + (1 - \lambda)x^2) = d'_0 + d^p_1 \lambda \) say. In the next interval \( \lambda_1 \leq \lambda \leq \lambda_2 \), \( \theta(\lambda) = c'_0 + c^2(\lambda x^1 + (1 - \lambda)x^2) = d'_0 + d^p_1 + d^p_2 \lambda \), etc. As \( f(x) \) is continuous, \( \theta(\lambda) \) is continuous, so at \( \lambda = \lambda_1 \), the two functions \( d'_0 + d^p_1 \lambda \), \( d'_0 + d^p_2 \lambda \) have the same value, and so on.
As \( f(x) \) is convex, \( \theta(\lambda) \) which is \( f(x) \) on the line segment \( L \) must also be convex. So from Result 2.2 we must have \( d_1^{p_1} < d_1^{p_2} < d_1^{p_3} < \ldots < d_1^{p_s} \). From this and the continuity of \( \theta(\lambda) \) it can be verified that \( \theta(\lambda) = d_0^{p_b} + d_1^{p_b} \lambda \geq d_0^{p} + d_1^{p} \tilde{x} \) for all \( p \in \{p_1, \ldots, p_s\} \), that is,

\[
f(\tilde{x}) = c_0^{pb} + c^{p_b} \tilde{x} \geq c_0^p + c^p \tilde{x} \quad \text{for all } p \in \{p_1, \ldots, p_s\}.
\]

By varying the points \( x^1, x^2 \), the same argument leads to the conclusion that

\[
f(\tilde{x}) = c_0^{pb} + c^{p_b} \tilde{x} \geq c_0^p + c^p \tilde{x} \quad \text{for all } p = 1 \text{ to } r.
\]

Since this holds for all points \( \tilde{x} \), \( f(x) \) satisfies (2.2). \( \square \)

The function \( f(x) \) defined by (2.2) is called the pointwise supremum function of the set of affine functions \( \{c_0^p + c^p x : p = 1, \ldots, r\} \). Theorem 2.5 shows that a PL function defined on \( \mathbb{R}^n \) is convex iff it is the pointwise supremum of a finite set of affine functions. In fact, in all applications where PL convex functions of two or more variables appear, they are usually seen in the form of pointwise supremum functions only. So, equations like (2.2) have become the standard way for defining PL convex functions.

In the same way, the PL function \( h(x) \) defined on \( \mathbb{R}^n \) is concave iff it is the pointwise infimum of a finite set of affine functions, that is, it is of the form \( h(x) = \min \{c_0^p + c^p x : p = 1 \text{ to } r\} \) for each \( x \in \mathbb{R}^n \).

In Fig. 2.6 we illustrate a pointwise supremum function \( \theta(\lambda) \) of a single variable \( \lambda \). \( \lambda \) is plotted on the horizontal axis, and the values of the function are plotted along the vertical axis. The function plotted is the pointwise supremum

![Fig. 2.6 Convexity and pointwise supremum property of a function of one variable. The various functions of which it is supremum are called \( a_1(\lambda) \) to \( a_4(\lambda) \).](image-url)
\[ \theta(\lambda) = \max\{a_1(\lambda) = 1 - 2\lambda, \ a_2(\lambda) = 1 + 0\lambda, \ a_3(\lambda) = -1 + \lambda, \ a_4(\lambda) = -4 + 2\lambda\}. \] The graph of \( \theta(\lambda) \) is plotted in the figure with thick lines. The function is:

<table>
<thead>
<tr>
<th>Interval</th>
<th>( \theta(\lambda) )</th>
<th>Slope in interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda \leq 0 )</td>
<td>( 1 - 2\lambda )</td>
<td>-2</td>
</tr>
<tr>
<td>( 0 \leq \lambda \leq 2 )</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>( 2 \leq \lambda \leq 3 )</td>
<td>( -1 + \lambda )</td>
<td>1</td>
</tr>
<tr>
<td>( \lambda \geq 3 )</td>
<td>( -4 + 2\lambda )</td>
<td>2</td>
</tr>
</tbody>
</table>

In Fig. 2.7, we illustrate a PL concave function \( h(\lambda) \) of a single variable \( \lambda \), which is the pointwise infimum \( h(\lambda) = \min\{a_1(\lambda) = 4 + \lambda, \ a_2(\lambda) = 3 + (1/2)\lambda, \ a_3(\lambda) = 3 - \lambda, \ a_4(\lambda) = 4 - 2\lambda\} \). The graph of \( h(\lambda) \) is shown in thick lines. This function is:

<table>
<thead>
<tr>
<th>Interval</th>
<th>( h(\lambda) )</th>
<th>Slope in interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda \leq -2 )</td>
<td>4 + ( \lambda )</td>
<td>1</td>
</tr>
<tr>
<td>( -2 \leq \lambda \leq 0 )</td>
<td>3 + (1/2}\lambda</td>
<td>1/2</td>
</tr>
<tr>
<td>( 0 \leq \lambda \leq 1 )</td>
<td>3 - ( \lambda )</td>
<td>-1</td>
</tr>
<tr>
<td>( \lambda \geq 1 )</td>
<td>4 - 2( \lambda )</td>
<td>-2</td>
</tr>
</tbody>
</table>

**Fig. 2.7** Concavity and pointwise infimum property of a function of one variable. The various functions of which it is infimum are called \( a_1(\lambda) \) to \( a_4(\lambda) \)
Exercises

2.3.6. Considering functions of decision variables \( x = (x_1, \ldots, x_n)^T \) defined over \( R^n \), prove that: (1) the sum of convex (concave) functions is convex (concave), (2) any positive combination of convex (concave) functions is convex (concave), (3) pointwise supremum of convex functions is convex, likewise pointwise infimum of concave functions is concave.

2.3.7. (1) Consider the function \( f(x) \) of a real-valued variable \( x \). Draw the graph of \( f(x) \) and show that it is a PL convex function. (2) In the same way show that \( f(x) = c|x| \), where \( c \) is a constant, is PL convex if \( c \geq 0 \), and PL concave if \( c \leq 0 \). (3) Draw the graphs of the absolute values of affine functions \(|4 + \lambda|\) and \(|4 - 2\lambda|\) and show that these functions are PL convex. (4) For any \( j = 1 \) to \( n \), show that the function \( f(x) = |x_j| \) of \( x = (x_1, \ldots, x_n)^T \) defined over \( R^n \) is PL convex. What are the regions of \( R^n \) within which it is linear? (5) Show that the function \( f(x) = \sum_{j=1}^{n} c_j |x_j| \) defined over \( R^n \) is convex if \( c_j \geq 0 \) for all \( j \), concave if \( c_j \leq 0 \) for all \( j \). (6) Show that the absolute value function \( f(x) = |c_0 + cx| \) of \( x \in R^n \) is convex. What are the regions of \( R^n \) within which it is linear? Express this function as the pointwise supremum of a set of affine functions. (7) Show that the function \( f(x) = \sum_{r=1}^{l} w_r |c_r^T x| \) (linear combinations of affine functions) is convex if \( w_r \geq 0 \) for all \( r \), concave if \( w_r \leq 0 \) for all \( r \).

2.3.8. Consider the real-valued continuous function \( f(\lambda) \) of a variable \( \lambda \), defined over \( \lambda \geq 0 \); with \( f(0) = -20 \); and slopes of 5, 9, 11, 8, 6, 10, respectively, in the intervals \( [0, 20] \), \( [20, 50] \), \( [50, 60] \), \( [60, 80] \), \( [80, 90] \), \( [90, \infty] \). Is it a convex or a concave function over \( R \geq 0 \)? If not, are there convex subsets of \( R \) on which this function is convex or concave? If so, mention these and explain the reasons for the same.

2.3.9. Consider a function \( \theta(x) \) defined over a convex set \( \Gamma \subset R^n \). A point \( \bar{x} \in \Gamma \) is said to be a local minimum for \( \theta(x) \) over \( \Gamma \) if \( \theta(x) \geq \theta(\bar{x}) \) for all points \( x \in \Gamma \) satisfying \(|x - \bar{x}| \leq \epsilon \) for some \( \epsilon > 0 \).

A local minimum \( \bar{x} \) for \( \theta(x) \) in \( \Gamma \) is said to be its global minimum in \( \Gamma \) if \( \theta(x) \geq \theta(\bar{x}) \) for all points \( x \in \Gamma \). Local maximum, global maximum have corresponding definitions.

Prove that every local minimum [maximum] of \( \theta(x) \) in \( \Gamma \) is a global minimum [maximum] if \( \theta(x) \) is convex [concave]. Construct simple examples of general functions defined over \( R \) which do not satisfy these properties.

Also, construct an example of a convex function that has a local maximum that is not a global maximum.

2.3.10. Show that the function \( f(\lambda) = |\lambda + 1| + |\lambda - 1| \) defined on \( R \) is convex, and that it has many local minima all of which are its global minima.
2.4 Optimizing PL Functions Subject to Linear Constraints

The problem of optimizing a general continuous PL function subject to linear constraints is a hard problem for which there are no known efficient algorithms. Some of these problems can be modeled as integer programs and solved by enumerative methods known for integer programs. These enumerative methods are fine for handling small-size problems, but require too much computer time as the problem size increases. However, the special problems of either:

- Minimizing a PL convex function, or equivalently
- Maximizing a PL concave function

subject to linear constraints can be transformed into LPs by introducing additional variables, and solved by efficient algorithms available for LPs. We will now discuss these transformations with several illustrative examples.

2.4.1 Minimizing a Separable PL Convex Function Subject to Linear Constraints

The negative of a concave function is convex. Maximizing a concave function is the same as minimizing its negative, which is a convex function. Using this, the techniques discussed here can also be used to solve problems in which a separable PL concave function is required to be maximized subject to linear constraints.

A real-valued function \( z(x) \) of decision variables \( x = (x_1, \ldots, x_n)^T \) is said to be a separable function if it can be expressed as the sum of \( n \) different functions, each one involving only one variable, that is, has the form \( z(x) = z_1(x_1) + z_2(x_2) + \ldots + z_n(x_n) \). This separable function is also a PL convex function if \( z_j(x_j) \) is a PL convex function for each \( j = 1 \) to \( n \).

**Result 2.3.** Let \( \theta(\lambda) \) be the PL convex function of \( \lambda \in \mathbb{R}^1 \) defined over \( \lambda \geq 0 \) shown in the following table:

<table>
<thead>
<tr>
<th>Interval</th>
<th>Slope</th>
<th>( \theta(\lambda) = )</th>
<th>Interval length</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 0 = \lambda_0 \leq \lambda \leq \lambda_1 )</td>
<td>( c_1 )</td>
<td>( c_1 \lambda )</td>
<td>( \lambda_1 )</td>
</tr>
<tr>
<td>( \lambda_1 \leq \lambda \leq \lambda_2 )</td>
<td>( c_2 )</td>
<td>( \theta(\lambda_1) + c_2(\lambda - \lambda_1) )</td>
<td>( \lambda_2 - \lambda_1 )</td>
</tr>
<tr>
<td>( \lambda_2 \leq \lambda \leq \lambda_3 )</td>
<td>( c_3 )</td>
<td>( \theta(\lambda_2) + c_3(\lambda - \lambda_2) )</td>
<td>( \lambda_3 - \lambda_2 )</td>
</tr>
<tr>
<td>[ \vdots ]</td>
<td>[ \vdots ]</td>
<td>[ \vdots ]</td>
<td>[ \vdots ]</td>
</tr>
<tr>
<td>( \lambda_{r-1} \leq \lambda \leq \lambda_r = \infty )</td>
<td>( c_r )</td>
<td>( \theta(\lambda_{r-1}) + c_r(\lambda - \lambda_{r-1}) )</td>
<td>( \infty )</td>
</tr>
</tbody>
</table>
where \( \lambda_1 < \lambda_2 < \ldots < \lambda_{r-1} \) and \( c_1 < c_2 < \ldots < c_r \) (conditions for \( \theta(\lambda) \) to be convex). Then for any \( \lambda \geq 0 \), \( \theta(\lambda) \) is the minimum objective value in the following problem.

Minimize \( z = c_1 \mu_1 + \ldots + c_r \mu_r \)
subject to \( \mu_1 + \ldots + \mu_r = \tilde{\lambda} \) \( (2.5) \)
\[ 0 \leq \mu_t \leq \lambda_t - \lambda_{t-1} \quad t = 1, \ldots, r \]

**Proof.** Problem (2.5) can be interpreted this way: Suppose we want to purchase exactly \( \tilde{\lambda} \) units of a commodity for which there are \( r \) suppliers. For \( k = 1 \) to \( r \), \( k \)th supplier’s rate is \( c_k/\text{unit} \) and can supply up to \( \lambda_k - \lambda_{k-1} \) units only. \( \mu_k \) in the problem represents the amount purchased from the \( k \)th supplier, it is \( 0 \), but is bounded above by the length of the \( k \)th interval in which the slope of \( \theta(\lambda) \) is \( c_k \). \( z \) to be minimized is the total expense to acquire the required \( \tilde{\lambda} \) of the commodity.

Clearly, to minimize \( z \), we should purchase as much as possible from the cheapest supplier, and when he cannot supply any more go to the next cheapest supplier, and continue the same way until the required quantity is acquired. As the cost coefficients satisfy \( c_1 < c_2 < \ldots < c_r \) by the convexity of \( \theta(\lambda) \), the cheapest cost coefficient corresponds to the leftmost interval beginning with \( 0, \) the next cheapest corresponds to the next interval just to the right of it, and so on. Because of this, the optimum solution \( \tilde{\mu} = (\tilde{\mu}_1, \ldots, \tilde{\mu}_r) \) of (2.5) satisfies the following special property.

**Special property of optimum solution \( \tilde{\mu} \) of (2.5) that follows from convexity of \( \theta(\lambda) \):** If \( p \) is such that \( \lambda_p \leq \tilde{\lambda} \leq \lambda_{p+1} \), then \( \tilde{\mu}_t = \lambda_t - \lambda_{t-1} \), the upper bound of \( \mu_t \) for all \( t = 1 \) to \( p \), \( \tilde{\mu}_{p+1} = \lambda - \lambda_p \), and \( \tilde{\mu}_t = 0 \) for all \( t \geq p + 2 \).

This property says that in the optimum solution of (2.5) if any \( \mu_k > 0 \), then the value of \( \mu_t \) in it must be equal to the upper bound on this variable for any \( t < k \). Because of this, the optimum objective value in (2.5) is \( z = c_1 \tilde{\mu}_1 + \ldots + c_r \tilde{\mu}_r \theta(\tilde{\lambda}) \).

**Example 2.3. – Illustration of Result 2.3:** Consider the following PL function.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Slope in interval</th>
<th>( \theta(\lambda) = )</th>
<th>Interval length</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–10</td>
<td>1</td>
<td>( \lambda )</td>
<td>10</td>
</tr>
<tr>
<td>10–25</td>
<td>2</td>
<td>( 10 + 2(\lambda - 10) )</td>
<td>15</td>
</tr>
<tr>
<td>25–30</td>
<td>4</td>
<td>( 40 + 4(\lambda - 25) )</td>
<td>5</td>
</tr>
<tr>
<td>30–( \infty )</td>
<td>6</td>
<td>( 60 + 6(\lambda - 30) )</td>
<td>( \infty )</td>
</tr>
</tbody>
</table>

As the slope is increasing with \( \lambda \), \( \theta(\lambda) \) is convex. Consider \( \tilde{\lambda} = 27 \). We see that \( \theta(27) = 48 \). The LP corresponding to (2.5) for \( \tilde{\lambda} = 27 \) in this problem is

Minimize \( z = \mu_1 + 2\mu_2 + 4\mu_3 + 6\mu_4 \)
subject to \( \mu_1 + \mu_2 + \mu_3 + \mu_4 = 27 \)
\[ 0 \leq \mu_1 \leq 10, \quad 0 \leq \mu_2 \leq 15 \]
\[ 0 \leq \mu_3 \leq 5, \quad 0 \leq \mu_4 \]
The optimum solution of this LP is obtained by increasing the values of \( \mu_1, \mu_2, \mu_3, \mu_4 \) one at a time from 0 in this order, moving to the next when this reaches its upper bound, until the sum of these variables reaches 27. So, the optimum solution is \( \bar{\mu} = (10, 15, 2, 0)^T \) with its objective value of \( 2\bar{\mu}_1 + 4\bar{\mu}_2 + 6\bar{\mu}_3 + 6\bar{\mu}_4 = 48 = \theta(27) \) computed earlier from the definition of this function, verifying Result 2.3 in this example.

If \( \theta(\lambda) \) is not convex, the optimum solution of (2.5) will not satisfy the special property described in the proof of Result 2.3.

Because of this result when \( \theta(\lambda) \) is PL convex, in minimizing a PL convex function in which \( \theta(\lambda) \) is one of the terms, we can linearize \( \theta(\lambda) \) by replacing \( \lambda \) by \( \sum_{t=1}^{r'} \mu_t \), where \( \mu_t \) is a new nonnegative variable corresponding to the \( t \)th interval in the definition of \( \theta(\lambda) \), bounded above by the length of this interval, and replacing \( \theta(\lambda) \) by \( \sum_{t=1}^{r'} c_t \mu_t \).

So now consider the problem

\[
\text{Minimize } z(x) = z_1(x_1) + \ldots + z_n(x_n)
\]

subject to \( Ax = b \) \hspace{1cm} (2.6)
\[x \geq 0,\]

where, for each \( j \), \( z_j(x_j) \) is a PL convex function defined on \( x_j \geq 0 \). Suppose the various slopes for \( z_j(x_j) \) are \( c_j^1 < c_j^2 < \ldots < c_j^r_j \) in that order with slopes changing at the values \( d_j^1 < d_j^2 < \ldots < d_j^{r_j+1} = \text{length of the } k \text{th interval in the definition of } z_j(x_j) \).

Then from this discussion, the LP formulation for (2.6) involving new variables \( x_j^k \) for \( k = 1 \) to \( r_j \), \( j = 1 \) to \( n \) is (here \( \ell_j^k = d_j^k - d_j^{k-1} \))

\[
\text{Minimize } \sum_{j=1}^{n} \sum_{k=1}^{r_j} c_j^k x_j^k
\]

subject to \( \sum_{k=1}^{r_j} x_j^k = x_j, \ j = 1 \) to \( n \)
\[Ax = b \]
\[x \geq 0 \]
\[0 \leq x_j^k \leq \ell_j^k, \hspace{0.5cm} 1 \leq j \leq n, \hspace{0.5cm} 1 \leq k \leq r_j \]

**Example 2.4.** A company makes products \( P_1, P_2, P_3 \) using limestone (L), electricity (E), water (W), fuel (F), and labor (L) as inputs. Labor is measured in man hours, other inputs in suitable units. Each input is available from one or more sources.
The company has its own quarry for LI, which can supply up to 250 units/day at a cost of $20/unit. Beyond that, LI can be purchased in any amounts from an outside supplier at $50/unit.

EP is available only from the local utility. Their charges for EP are $30/unit for the first 1,000 units/day, $45/unit for up to an additional 500 units/day beyond the initial 1,000 units/day, $75/unit for amounts beyond 1,500 units/day.

Up to 800 units/day of W (water) is available from the local utility at $6/unit, beyond that they charge $7/unit of water/day.

There is a single supplier for F who can supply at most 3,000 units/day at $40/unit, beyond that there is currently no supplier for F.

From their regular workforce they have up to 640 man hours of labor/day at $10/man hour, beyond that they can get up to 160 man hours/day at $17/man hour from a pool of workers.

They can sell up to 50 units of \( P_1 \) at $3,000/unit/day in an upscale market; beyond that they can sell up to 50 more units/day of \( P_1 \) to a wholesaler at $250/unit. They can sell up to 100 units/day of \( P_2 \) at $3,500/unit. They can sell any quantity of \( P_3 \) produced at a constant rate of $4,500/unit.

Data on the inputs needed to make the various products is given in the following table. Formulate the product mix problem to maximize the net profit/day at this company.

<table>
<thead>
<tr>
<th>Product</th>
<th>Input units/unit made</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LI</td>
</tr>
<tr>
<td>( P_1 )</td>
<td>1/2</td>
</tr>
<tr>
<td>( P_2 )</td>
<td>1</td>
</tr>
<tr>
<td>( P_3 )</td>
<td>3/2</td>
</tr>
</tbody>
</table>

Maximizing the net profit is the same thing as minimizing its negative, which is \( = (\text{the costs of all the inputs used/day}) - (\text{sales revenue/day}) \). We verify that each term in this sum is a PL convex function. So, we can model this problem as an LP in terms of variables corresponding to each interval of constant slope of each of the input and output quantities.

Let LI, EP, W, F, L denote the quantities of the respective inputs used/day; and \( P_1, P_2, P_3 \) denote the quantities of the respective products made and sold/day. Let LI\(_1, \) LI\(_2 \) denote units of limestone used daily from own quarry, outside supplier. Let EP\(_1, \) EP\(_2, \) EP\(_3 \) denote units of electricity used/day at $30, 45, 75/unit, respectively. Let W\(_1, \) W\(_2 \) denote units of water used/day at rates of $6 and 7/unit, respectively. Let L\(_1, \) L\(_2 \) denote the man hours of labor used/day from regular workforce, pool, respectively. Let \( P_{11}, P_{12} \) denote the units of \( P_1 \) sold at the upscale market, to the wholesaler, respectively.

Then the LP model for the problem is

\[
\text{Minimize } z = 20LI_1 + 50LI_2 + 30EP_1 + 45EP_2 + 75EP_3 + 6W_1 + 7W_2 + 40F \\
+ 10L_1 + 17L_2 - 3,000P_{11} - 250P_{12} - 3,500P_2 - 4,500P_3
\]
subject to

\[
\begin{align*}
(1/2) P_1 + P_2 + (3/2) P_3 &= LI \\
3P_1 + 2P_2 + 5P_3 &= EP \\
P_1 + (1/4)P_2 + 2P_3 &= W \\
P_1 + P_2 + 3P_3 &= F \\
2P_1 + P_2 + P_3 &= L \\
L_1 + L_2 &= L, \\
W_1 + W_2 &= W \\
L_1 + L_2 &= L, \\
P_{11} + P_{12} &= P_1, & \text{All variables} \geq 0
\end{align*}
\]

(\(LI_1, EP_1, EP_2, W_1\)) \leq (250, 1,000, 500, 800)
(\(F, L_1, L_2\)) \leq (3,000, 640, 160)
(\(P_{11}, P_{12}, P_2\)) \leq (50, 50, 100). 

### 2.4.2 Min-max, Max-min Problems

As discussed earlier, a PL convex function in variables \(x = (x_1, \ldots, x_n)^T\) can be expressed as the pointwise maximum of a finite set of affine functions. Minimizing a function like that subject to some constraints is appropriately known as a min-max problem.

Similarly, a PL concave function in \(x\) can be expressed as the pointwise minimum of a finite set of affine functions. Maximizing a function like that subject to some constraints is appropriately known as a max-min problem. Both min-max and max-min problems can be expressed as LPs using just one additional variable, if all the constraints are linear constraints.

If the PL convex function \(f(x) = \max \{c_0^t + c^t x : t = 1, \ldots, r\}\), then \(-f(x) = \min \{-c_0^t - c^t x : t = 1, \ldots, r\}\) is PL concave and conversely. Using this, any min-max problem can be posed as a max-min problem and vice versa. So, it is sufficient to discuss max-min problems. Consider the max-min problem

\[
\begin{align*}
\text{Maximize } z(x) &= \min \{c_0^t + c^t x, \ldots, c_0^r + c^r x\} \\
\text{subject to } Ax &= b \\
x \geq 0.
\end{align*}
\]

To transform this problem into an LP, introduce the new variable \(x_{n+1}\) to denote the value of the objective function \(z(x)\) to be maximized. Then the equivalent LP with additional linear constraints is

\[
\begin{align*}
\text{Maximize } & x_{n+1} \\
\text{subject to } & x_{n+1} \leq c_0^1 + c^1 x \\
& x_{n+1} \leq c_0^2 + c^2 x \\
& \vdots
\end{align*}
\]
The fact that $x_{n+1}$ is being maximized and the additional constraints together imply that if $(\bar{x}, \bar{x}_{n+1})$ is an optimum solution of this LP model, then $ar{x}_{n+1} = \min\{c_0^T + c^T \bar{x}, \ldots, c_0^T + c^T \bar{x}\} = z(\bar{x})$, and that $\bar{x}_{n+1}$ is the maximum value of $z(x)$ in the original max-min problem.

Example 2.5. **Application of the Min-max Model in Worst Case Analysis:** Consider the fertilizer maker’s product mix problem with decision variables $x_1, x_2$ (hi-ph, lo-ph fertilizers to be made daily in the next period) discussed in Sect. 1.7.1 and in Example 3.4.1 of Sect. 3.4 of Murty (2005b) of Chap. 1. This company makes hi-ph, lo-ph fertilizers using raw materials RM1, RM2, RM3 with the following data (Table 2.1):

We discussed the case where the net profit coefficients $c_1, c_2$ of these variables are estimated to be $15$ and $10$, respectively. In reality, the prices of fertilizers are random variables that fluctuate daily. Because of unstable conditions and new agricultural research announcements, suppose that market analysts have only been able to estimate that the expected net profit coefficient vector $(c_1, c_2)$ is likely to be one of $\{(15, 10), (10, 15), (12, 12)\}$ without giving a single point estimate. So, here we have three possible scenarios. In scenario 1, $(c_1, c_2) = (15, 10)$, expected net profit $= 15x_1 + 10x_2$; in scenario 2, $(c_1, c_2) = (10, 15)$, expected net profit $= 10x_1 + 15x_2$; and in scenario 3, $(c_1, c_2) = (12, 12)$, expected net profit $= 12x_1 + 12x_2$. Suppose the raw material availability data in the problem is expected to remain unchanged. The important question is: which objective function to optimize for determining the production plan for next period.

Regardless of which of the three possible scenarios materializes, at the worst the minimum expected net profit of the company will be $p(x) = \min\{15x_1 + 10x_2, 10x_1 + 15x_2, 12x_1 + 12x_2\}$ under the production plan $x = (x_1, x_2)^T$. **Worst case analysis** is an approach that advocates determining the production plan to optimize this worst case net profit $p(x)$ in this situation. This leads to the max-min model:

Maximize $p(x) = \min\{15x_1 + 10x_2, 10x_1 + 15x_2, 12x_1 + 12x_2\}$

<table>
<thead>
<tr>
<th>Table 2.1 Data for the fertilizer problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>Hi-ph</td>
</tr>
<tr>
<td>RM 1</td>
</tr>
<tr>
<td>RM 2</td>
</tr>
<tr>
<td>RM 3</td>
</tr>
<tr>
<td>Net profit $/ton made</td>
</tr>
</tbody>
</table>
subject to $2x_1 + x_2 \leq 1.500$
$x_1 + x_2 \leq 1.200$
$x_1 \leq 500$
$x_1, x_2 \geq 0$.

Its LP formulation is

$$
\begin{align*}
\text{max} & \quad p \\
\text{subject to} & \quad p \leq 15x_1 + 10x_2 \\
& \quad p \leq 10x_1 + 15x_2 \\
& \quad p \leq 12x_1 + 12x_2 \\
& \quad 2x_1 + x_2 \leq 1.500 \\
& \quad x_1 + x_2 \leq 1.200 \\
& \quad x_1 \leq 500, \quad x_1, x_2 \geq 0.
\end{align*}
$$

2.4.3 Minimizing Positive Linear Combinations of Absolute Values of Affine Functions

Let $z(x) = w_1 |c_0^1 + c^1 x| + \ldots + w_r |c_0^r + c^r x|$. Consider the problem:

$$
\begin{align*}
\text{Minimize} & \quad z(x) \\
\text{subject to} & \quad Ax \geq b,
\end{align*}
$$

(2.8)

where the weights $w_1, \ldots, w_r$ are all strictly positive. In this problem the objective function to be minimized, $z(x)$, is a PL convex function, hence this problem can be transformed into an LP. This is based on a result that helps to express the absolute value as a linear function of two additional variables, which we will discuss first.

Result 2.4. Consider the affine function $c_0^k + c^k x$ and its value $\beta = c_0^k + c^k \bar{x}$ at some point $\bar{x} \in \mathbb{R}^n$. Consider the following LP in two variables $u, v$.

$$
\begin{align*}
\text{Minimize} & \quad u + v \\
\text{subject to} & \quad u - v = \beta \\
& \quad u, v \geq 0
\end{align*}
$$

(2.9)

(2.9) has a unique optimum solution $(\bar{u}, \bar{v})$, which satisfies $\bar{u} \bar{v} = 0$, and its optimum objective value $\bar{u} + \bar{v} = |\beta| = |c_0^k + c^k \bar{x}|$. 
Proof. If $\beta \geq 0$, the general solution of (2.9) is $(u, v) = (\beta + \alpha, \alpha)$ for some $\alpha \geq 0$, the objective value of this solution, $\beta + 2\alpha$, assumes its minimum value when $\alpha = 0$. So in this case $(\bar{u}, \bar{v}) = (\beta, 0)$ satisfying $\bar{u}\bar{v} = 0$ and having optimum objective value of $\bar{u} + \bar{v} = \beta = |\beta|$.

If $\beta < 0$, the general solution of (2.9) is $(u, v) = (\alpha, |\beta| + \alpha)$ for some $\alpha \geq 0$, the objective value of this solution, $|\beta| + 2\alpha$, assumes its minimum value when $\alpha = 0$. So in this case $(\bar{u}, \bar{v}) = (0, |\beta|)$ satisfying $\bar{u}\bar{v} = 0$ and having optimum objective value of $\bar{u} + \bar{v} = |\beta|$.

So, the result holds in all cases. □

Example 2.6. Illustration of Result 2.4: Consider problem (2.9) when $\beta = -7$. The problem is

minimize $u + v$ subject to $u - v = -7, u, v \geq 0$.

The general solution of this problem is $(u, v) = (\alpha, 7 + \alpha)$ for $\alpha \geq 0$ with objective value $7 + 2\alpha$. So, the unique optimum solution is $(\bar{u}, \bar{v}) = (0, 7)$ and

$$\bar{u} + \bar{v} = 7 = |-7| \text{ and } \bar{u}\bar{v} = 0.$$ 

In the optimum solution $(\bar{u}, \bar{v})$ of (2.9), $\bar{u}$ is usually called the positive part of $\beta$, and $\bar{v}$ is called the negative part of $\beta$. Notice that when $\beta$ is negative, its negative part is actually the absolute value of $\beta$. Also, for all values of $\beta$, at least one quantity in the pair (positive part of $\beta$, negative part of $\beta$) is 0.

Commonly the positive or negative parts of $\beta$ are denoted by symbols $\beta^+, \beta^-$, respectively. In this notation, $\beta = \beta^+ - \beta^-$ and $|\beta| = \beta^+ + \beta^-$; both $\beta^+, \beta^-$ are $\geq 0$, and satisfy $(\beta^+)(\beta^-) = 0$.

Result 2.4 helps to linearize the objective function in (2.8) by introducing two new variables for each absolute value term in it. Notice that this is only possible when all the coefficients of the absolute value terms in the objective function in (2.8) are positive. From this discussion we see that (2.8) is equivalent to the following LP with two new nonnegative variables for each $t = 1$ to $r$, $u_t^+ = \max \{0, c_0^t + c^t x\}$, $u_t^- = -\min \{0, c_0^t + c^t x\}$. $u_t^+$ is the positive part of $c_0^t + c^t x$ and $u_t^-$ its negative part.

Minimize $w_1[(u_1^+) + (u_1^-)] + \ldots + w_r[(u_r^+) + (u_r^-)]$

subject to $c_0^1 + c^1 x = (u_1^+) - (u_1^-)$

$$\vdots$$

$c_0^r + c^r x = (u_r^+) - (u_r^-)$

$Ax \geq b$

$(u_t^+), (u_t^-) \geq 0, \quad t = 1, \ldots, r.$

If $(\hat{u}_1^+, \ldots, \hat{u}_r^+), (\hat{u}_1^-, \ldots, \hat{u}_r^-), (\hat{x})$ is an optimum solution of (2.10), then $\hat{x}$ is an optimum solution of (2.8), and $c_0^k + c^k \hat{x} = \hat{u}_k^+ - \hat{u}_k^-$, $|c_0^k + c^k \hat{x}| = \hat{u}_k^+ + \hat{u}_k^-$; and the optimum objective values in (2.10) and (2.8) are the same.
Application of this transformation will be discussed next. This is an important model that finds many applications.

In Model (2.10), by expressing the affine function $c_0^1 + c^1 x$, which may be positive or negative, as the difference $u^+ - u^-$ of two nonnegative variables; the positive part of $c_0^1 + c^1 x$ denoted by $(c_0^1 + c^1 x)^+ = \max\{c_0^1 + c^1 x, 0\}$ will be $u^+$, and the negative part of $c_0^1 + c^1 x$ denoted by $(c_0^1 + c^1 x)^- = \max\{0, -(c_0^1 + c^1 x)\}$ will be $u^-$ as long as the condition $(u^+)(u^-) = 0$ holds. This condition will automatically hold as long as:

1. The coefficients of $u^+$, $u^-$ are both $\geq 0$ in the objective function being minimized; and
2. The column vectors of the pair of variables $u^+$, $u^-$ in the model among the constraints (not including the sign restrictions) sum to 0 (or form a linearly dependent set).

A Cautionary Note 2.1: When expressing an unrestricted variable or an affine function as a difference $u^+ - u^-$ of two nonnegative variables, and using $u^+$, $u^-$ as the positive, negative parts of that unrestricted variable or affine function, or using $|u|^\Delta$ as its absolute value, it is necessary to make sure that the condition $(u^+)(u^-) = 0$ will automatically hold at very optimum solution of the model. For this, the above two conditions must hold.

Sometimes people tend to include additional constraints involving $u^+$, $u^-$ with nonzero coefficients into the model (for examples, see Model 1 below, and Model 1 for the parameter estimation problem using the $L_\infty$-measure of deviation in Example 2.8 below). When this is done, the Condition 2 above may be violated; this may result in the model being invalid. So, it is better to not include additional constraints involving $u^+$, $u^-$ into the model.

2.4.4 Minimizing the Maximum of the Absolute Values of Several Affine Functions

Let $z(x) = \max\{|c_0^1 + c^1 x|, \ldots, |c_r^0 + c^r x|\}$. Consider the problem

\[ \text{Minimize } \quad z(x) \]
\[ \text{subject to } \quad Ax \geq b. \]  
(2.11)

In this problem the objective function to be minimized, $z(x)$, is the pointwise supremum of several PL convex functions, and hence is a PL convex function, hence this problem can be transformed into an LP. Combining the ideas discussed above, one LP model for this problem is Model 1 given below.

It can be verified that in this model the property $(u^+_t)(u^-_t) = 0$ for all $t$ will hold at every optimum solution for it, so this is a valid model for the problem. But it has one disadvantage that it uses the variables $u^+_t$, $u^-_t$ representing the positive and negative parts of $c^t_0 + c^t x$ in additional constraints in the model (those in the first line of constraints), with the result that the pair of column vectors of the variables
Among the constraints no longer form a linearly dependent set, violating Condition 2 expressed in Cautionary Note 2.1 above.

**Model 1**

\[
\begin{align*}
\min \quad & z \\
\text{subject to} \quad & z \geq u^+_t + u^-_t, \quad t = 1,\ldots,r \\
& c_0^1 + c^1 x = u^+_1 - u^-_1 \\
& \quad \vdots \\
& c_0^r + c^r x = u^+_r - u^-_r \\
& Ax \geq b \\
& u^+_t, u^-_t \geq 0, \quad t = 1,\ldots,r
\end{align*}
\]  

(2.12)

It is possible to transform (2.11) into an LP model directly without introducing these \(u^+_t, u^-_t\) variables at all. This leads to a better and cleaner LP model for this problem, Model 2, with only one additional variable \(z\).

**Model 2**

\[
\begin{align*}
\min \quad & z \\
\text{subject to} \quad & -z \leq c_0^t + c^t x \leq z, \quad t = 1,\ldots,r \\
& Ax \geq b \\
& z \geq 0.
\end{align*}
\]  

(2.13)

The constraints specify that \(z \geq |c_0^t + c^t x|\) for all \(t\); and as \(z\) is minimized in Model 2, it guarantees that if \((\hat{z}, \hat{x})\) is an optimum solution of this Model 2, then \(\hat{x}\) is an optimum solution also for (2.11), and \(\hat{z}\) is the optimum objective value in (2.11).

We will now discuss important applications of these transformations in meeting multiple targets as closely as possible, and in curve fitting, and provide simple numerical examples for each.

**Example 2.7. Meeting targets as closely as possible:** Consider the fertilizer maker’s product mix problem with decision variables \(x_1, x_2\) (hi-ph, lo-ph fertilizers to be made daily in the next period) discussed in Example 3.4.1 of Sect. 3.4 of Murty (2005b) of Chap. 1 and Example 2.5 above, with net profit coefficients \((c_1, c_2) = (15, 10)\) in $/ton of hi-ph, lo-ph fertilizers made. In these examples, we considered only maximizing one objective function, the daily net profit \(= 15x_1 + 10x_2\) with the profit vector given. But in real business applications, companies have to pay attention to many other objective functions in order to survive and thrive in the market place. We will consider two others.

The second objective function that we will consider is the company’s total market share, usually measured by the company’s sales volume as a percentage of the sales volume of the whole market. To keep this example simple, we will measure this by
the total daily sales revenue of the company. The sale prices of hi-ph, lo-ph fertilizers are $222, $107/ton, respectively, so this objective function is $222x_1 + 107x_2$.

The third objective function that we consider is the hi-tech market share, which is the market share of the company among hi-tech products (in this case hi-ph is the hi-tech product). This influences the public’s perception of the company as a market leader. To keep this example simple, we will measure this by the daily sales revenue of the company from hi-ph sales which is $222x_1$.

So, here we have three different objective functions to optimize simultaneously. Problems like this are called multiobjective optimization problems. One commonly used technique to get a good solution in these problems is to set up a target value for each objective function (based on the companies aspirations, considering the trade-offs between the various objective functions), and to try to find a solution as close to each of the targets as possible. In our example, suppose that the targets selected for daily net profit, market share, and hi-tech market share are $12,500, 200,000, and 70,000, respectively.

In this example, we consider the situation where the company wants to attain the target value for each objective function as closely as possible, considering both positive and negative deviations from the targets as undesirable.

When there is more than one objective function to be optimized simultaneously, decision makers may not consider all of them to be of the same importance. To account for this, it is customary to specify positive weights corresponding to the various objective functions, reflecting their importance, with the understanding that the higher the weight the more important it is to keep the deviation in the value of this objective function from its target small. So, this weight for an objective function plays the role of a penalty for unit deviation in this objective value from its target. In our example, suppose these weights for daily net profit, market share, and hi-tech market share, are 10, 6, and 8, respectively.

After these weights are given, one strategy to solve this problem is to determine the solution to implement to minimize the penalty function, which is the weighted sum of absolute deviations from the targets. This problem is (constraints on the decision variables are given in Example 2.5 above)

\[
\text{Minimize penalty function} = 10|15x_1 + 10x_2 - 12,500| + 6|222x_1 + 107x_2 - 200,000| + 8|222x_1 - 70,000|
\]

subject to \[2x_1 + x_2 \leq 1,500\]

\[x_1 + x_2 \leq 1,200\]

\[x_1 \leq 500\]

\[x_1, x_2 \geq 0.\]

Linearizing this leads to the following LP:

\[
\text{Minimize penalty function} = 10(u_1^+ - u_1^-) + 6(u_2^+ + u_2^-) + 8(u_3^+ + u_3^-)
\]
subject to  \[ 15x_1 + 10x_2 - 12,500 = u_1^+ - u_1^- \]
\[ 222x_1 + 107x_2 - 200,000 = u_2^+ - u_2^- \]
\[ 222x_1 + 70,000 = u_3^+ - u_3^- \]
\[ 2x_1 + x_2 \leq 1,500 \]
\[ x_1 + x_2 \leq 1,200 \]
\[ x_1 \leq 500 \]
\[ x_1, x_2, u_i^+, u_i^- \geq 0, \text{ for all } i. \]

If \( \hat{u}^+ = (\hat{u}_1^+, \hat{u}_2^+, \hat{u}_3^+) \), \( \hat{u}^- = (\hat{u}_1^-, \hat{u}_2^-, \hat{u}_3^-) \), \( \hat{x} = (\hat{x}_1, \hat{x}_2) \) is an optimum solution of this LP, then \( \hat{x} \) is an optimum solution that minimizes the penalty function.

**Example 2.8.** Best \( L_1 \) or \( L_\infty \) Approximations for Parameter Estimation in Curve Fitting Problems:

A central problem in science and technological research is to determine the optimum operating conditions of processes to maximize the yield from them. Let \( y \) denote the yield from a process whose performance is influenced by \( n \) controllable factors. Let \( x = (x_1, \ldots, x_n)^T \) denote the vector of values of these factors, and this vector characterizes how the process is run. So, here \( x = (x_1, \ldots, x_n)^T \) are the independent variables whose values the decision maker can control, and the yield \( y \) is the dependent variable whose value depends on \( x \). To model the problem of determining the optimum \( x \) mathematically, it is helpful to approximate \( y \) by a mathematical function of \( x \), which we will denote by \( y(x) \).

The data for determining the functional form of \( y(x) \) is the yield at several points \( x \in \mathbb{R}^n \) in the feasible range. As there are usually errors in the measurement of yield, one makes several measurement observations of the yield at each point \( x \) used in the experiment, and takes the average of these observations as the yield value at that point. The problem of determining the functional form of \( y(x) \) from such observed data is known as a curve fitting problem.

For a numerical example, consider the data in the following Table 2.2 obtained from experiments for the yield in a chemical reaction, as a function of the temperature \( t \) at which the reaction takes place.

The problem in this example is to determine a mathematical function \( y(t) \) that fits the observed data as closely as possible.

<table>
<thead>
<tr>
<th>Temperature ( t )</th>
<th>Yield, ( y(t) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-5</td>
<td>80</td>
</tr>
<tr>
<td>-3</td>
<td>92</td>
</tr>
<tr>
<td>-1</td>
<td>96</td>
</tr>
<tr>
<td>0</td>
<td>98</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
</tr>
</tbody>
</table>
The commonly used strategy to solve the curve-fitting problem for the dependent variable, yield \( y(x) \), in terms of independent variables, \( x = (x_1, \ldots, x_n)^T \), involves the following steps.

**Step 1: Model function selection:** Select a specific mathematical functional form \( f(x, a) \) with unknown parameters say \( a = (a_0, \ldots, a_k) \) (these parameters are things like coefficients of various terms, exponents, etc.) that seems to offer the best fit for the yield \( y(x) \).

In some cases there may be well-developed mathematical theory that specifies \( f(x, a) \) directly. If that is not the case, plots of \( y(x) \) against \( x \) can give an idea of suitable model functions to select.

For example, if plots indicate that \( y(x) \) appears to be linear in \( x \), then we can select the model function to be \( f(x, a) = a_0 + a_1 x_1 + \ldots + a_n x_n \), in which the coefficients \( a_0, a_1, \ldots, a_n \) are the unknown parameters. This linear model function is the most commonly used one in statistical theory, and the area of this theory that deals with determining the best values for these parameters by the method of least squares is called *linear regression theory*.

If plots indicate that \( y(x) \) appears to be quadratic in \( x \), then the model function to use is \( \sum_{i=1}^{n} \sum_{j=i}^{n} a_{ij} x_i x_j \) (where the coefficients \( a_{ij} \) are the parameters). Similarly, a cubic function in \( x \) may be considered as the model function if that appears more appropriate.

The linear, quadratic, cubic functions in \( x \) are special cases of the general polynomial function in \( x \). Selecting a polynomial function in \( x \) as the model function confers a special advantage for determining the best values for the unknown parameters because this model function is linear in these parameters.

When the number of independent variables \( n \) is not small (i.e., \( \geq 4 \)), using a complete polynomial function in \( x \) of degree \( \geq 2 \) as the model function leads to many unknown parameter values to be determined. That is why when such model functions are used, one normally uses the practical knowledge about the problem and the associated process to fix as many as possible of these unknown coefficients that are known to be insignificant with reasonable certainty at 0.

Polynomial functions of \( x \) of degree \( \leq 3 \) are the most commonly used model functions for curve-fitting. Functions outside this class are only used when there is supporting theory that indicates that they are more appropriate.

**Step 2: Selecting a measure of deviation:** Let \( f(x, a) \) be the model function selected to represent the yield, with \( a \) as the vector of parameters in it. Suppose the data available consists of \( r \) observations on the yield as in the following table.

<table>
<thead>
<tr>
<th>Independent vars.</th>
<th>( x^1 )</th>
<th>( x^2 )</th>
<th>( \ldots )</th>
<th>( x^r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed yield</td>
<td>( y^1 )</td>
<td>( y^2 )</td>
<td>( \ldots )</td>
<td>( y^r )</td>
</tr>
</tbody>
</table>

Then the deviations of the model function value from the observed yield at the data points \( x^1, \ldots, x^r \) are \( f(x^1, a) - y^1, \ldots, f(x^r, a) - y^r \). Some of these deviations may be \( \geq 0 \) and some \( \leq 0 \), but \( f(x, a) \) is considered to be a good fit for
the yield if all these deviations are small, that is, close to 0. In this step we have to select a single numerical measure that can check whether all these deviations are small or not.

The most celebrated and most commonly used measure of deviation is the sum of squared deviations, first used and developed by Carl F. Gauss, the famous nineteenth century German mathematician. He developed this measure for approximating the orbit of the asteroid Ceres with a second degree curve. This measure is also known as the $L_2$-measure (after the Euclidean or the $L_2$-metric defined as the square root of the sum of squares), and for our problem it is $L_2(a) = \sum_{k=1}^{r} (f(x^k, a) - y^k)^2$. Determining the best values of the parameters $a$ as those that minimizes this $L_2$ measure $L_2(a)$ is known as the method of least squares.

Another measure of deviation that can be used is the $L_1$-measure (also known as the rectilinear measure); it is the sum of absolute deviations $L_1(a) = \sum_{k=1}^{r} |f(x^k, a) - y^k|$.

A third measure of deviation that is used by some people is the $L_\infty$-measure (also known as the Chebyshev measure after the Russian mathematician Tschebychev who proposed it in the nineteenth century). This measure is the maximum absolute deviation $L_\infty(a) = \max\{ |f(x^k, a) - y^k| : k = 1 \text{ to } r \}$.

The $L_2$-measure is continuously differentiable in the parameters, but the $L_1$ and $L_\infty$-measures are not (they are not differentiable at points in the parameter space where a deviation term becomes 0). That is why minimizing the $L_2$-measure using calculus techniques based on derivatives is easier; for this reason the method of least squares has become a very popular method for determining the best values for the unknown parameters to give the best fit to the observed data. Particularly, most of statistical theory is based on the method of least squares.

As they are not differentiable at some points, minimizing the $L_1$ and $L_\infty$-measures may be difficult in general. However, when the model function $f(x, a)$ is linear in the parameter vector $a$ (this is the case when $f(x, a)$ is a polynomial in $x$), then determining $a$ to minimize the $L_1$ or $L_\infty$-measures can be transformed into LPs and solved very efficiently. That is why parameter estimation to minimize the $L_1$ or $L_\infty$-measures is becoming increasingly popular when $f(x, a)$ is linear in $a$.

The parameter vector that minimizes the $L_2$-measure is always unique, but the problem of minimizing $L_1$ or $L_\infty$-measures usually have alternate optima. There are some other differences among the $L_2$, $L_1$, $L_\infty$-measures worth noting. Many people do not like to use the $L_\infty$-measure for parameter estimation, because it determines the parameter values to minimize the deviations of extreme measurements (which are often labeled as “outliers” in statistical literature), totally ignoring all other observations. Both $L_1$, $L_2$-measures give equal weight to all the observations. The $L_2$-measure would be the preferred measure to use when $f(x, a)$ is not linear in the parameter vector $a$, because it is differentiable everywhere. When $f(x, a)$ is linear in $a$, the choice between $L_2$, $L_1$-measures of deviation to use for parameter estimation is a matter for individual judgement and the availability of suitable software for carrying out the computations required.
Step 3: Parameter estimation: Solve the problem of determining \( \tilde{a} \) that minimizes the measure of deviation selected.

The optimum solutions for the problems of minimizing \( L_2(a) \), \( L_1(a) \), \( L_\infty(a) \) may be different. Let \( \tilde{a} \) denote the optimum \( a \)-vector that minimizes whichever measure of deviation has been selected for determining the best \( a \)-vector. The optimum objective value in this problem is known as the residue. If the residue is “small,” \( f(x, \tilde{a}) \) is accepted as the functional form for \( y(x) \).

If the residue is “large,” it is an indication that \( f(x, a) \) is not the appropriate functional form for the yield \( y(x) \). In this case go back to Step 1 to select a better model function for the yield, and repeat this whole process with it.

Finally, the question of how to judge whether the residue is “small” or “large”. Statistical theory provides some tests of significance for this judgement when using the method of least squares. These are developed under the assumption that the observed yield follows a normal distribution. But, in general, the answer to this question depends mostly on personal judgement.

When \( f(x, a) \) is linear in \( a \), a necessary and sufficient condition for optimality for the problem of minimizing \( L_2(a) \) is \( \frac{\partial L_2(a)}{\partial a} = 0 \). This is a system of linear equations in \( a \), which can be solved for determining the optimum solution \( \tilde{a} \).

The problems of minimizing \( L_1(a) \) are \( L_\infty(a) \) when \( f(x, a) \) linear in \( a \) can be transformed into an LP. We will show how to do this using the example of yield in the chemical reaction as a function of the temperature \( t \) of the reaction; data for which is given in Table 2.2 above.

Estimates of the Parameter Vector \( a \) that Minimize \( L_2(a) \): Suppose plots indicate that the yield in this chemical reaction, as a function of the reaction temperature, \( y(t) \) can be approximated closely by a quadratic function of \( t \). So we take the model function to be \( f(t, a) = a_0 + a_1 t + a_2 t^2 \), where \( a = (a_0, a_1, a_2) \) is the parameter vector to be estimated.

So, \( f(-5, a) = a_0 - 5a_1 + 25a_2 \), hence the deviation between \( f(t, a) \) and \( y(t) \) at \( t = -5 \) is \( a_0 - 5a_1 + 25a_2 = 80 \). Continuing this way, we see that

\[
L_2(a) = (a_0 - 5a_1 + 25a_2 - 80)^2 + (a_0 - 3a_1 + 9a_2 - 92)^2 + (a_0 - a_1 + a_2 - 96)^2 \\
+ (a_0 - 98)^2 + (a_0 + a_1 + a_2 - 100)^2.
\]

\[
L_1(a) = |(a_0 - 5a_1 + 25a_2 - 80)| + |(a_0 - 3a_1 + 9a_2 - 92)| + |(a_0 - a_1 + a_2 - 96)| \\
+ |(a_0 - 98)| + |(a_0 + a_1 + a_2 - 100)|.
\]

\[
L_\infty\{(a) = \max\{|(a_0 - 5a_1 + 25a_2 - 80)|, |(a_0 - 3a_1 + 9a_2 - 92)| \\
| (a_0 - a_1 + a_2 - 96)|, |(a_0 - 98)|, |(a_0 + a_1 + a_2 - 100)|\}.
\]

So, the method of least squares involves finding \( a \) that minimizes \( L_2(a) \). The necessary and sufficient optimality conditions for this are \( \frac{\partial L_2(a)}{\partial a} = 0 \), which are

\[
5a_0 - 8a_1 + 36a_2 = 466,
-8a_0 + 36a_1 - 152a_2 = -672,
36a_0 - 152a_1 + 708a_2 = 3,024.
\]
It can be verified that this has the unique solution of \( \bar{a} = (\bar{a}_0, \bar{a}_1, \bar{a}_2) = (98.6141, 1.1770, -0.4904) \). So the fit obtained by the method of least squares is 
\[
f(t, \bar{a}) = 98.6141 + 1.1770t - 0.4904t^2,
\]
with a residue of 3.7527, in \( L_2 \)-measure units.

**Estimates of the Parameter Vector \( \mathbf{a} \) that Minimize \( L_1(\mathbf{a}) \):** The problem of minimizing \( L_1(\mathbf{a}) \) is the following LP:

\[
\text{Minimize } \sum_{i=1}^{5} (u_i^+ + u_i^-) \\
\text{subject to } (a_0 - 5a_1 + 25a_2 - 80) = u_1^+ - u_1^- \\
(a_0 - 3a_1 + 9a_2 - 92) = u_2^+ - u_2^- \\
(a_0 - a_1 + a_2 - 96) = u_3^+ - u_3^- \\
(a_0 - 98) = u_4^+ - u_4^- \\
(a_0 + a_1 + a_2 - 100) = u_5^+ - u_5^- \\
u_i^+, u_i^- \geq 0, \text{ for all } i.
\]

One of the optimum solutions of this problem is \( \hat{a} = (\hat{a}_0, \hat{a}_1, \hat{a}_2) = (98.3333, 2, -0.3333) \); the fit given by this solution is 
\[
f(t, \hat{a}) = 98.3333 + 2t - 0.3333t^2,
\]
with a residue of 3, in \( L_1 \)-measure units.

**Estimates of the Parameter Vector \( \mathbf{a} \) that Minimize \( L_\infty(\mathbf{a}) \):** One LP model discussed earlier for the problem of minimizing \( L_\infty(\mathbf{a}) \) is the following:

**Model 1:**

\[
\text{Minimize } z \\
\text{subject to } z \geq (u_i^+ + u_i^-) \text{ for all } i \\
(a_0 - 5a_1 + 25a_2 - 80) = u_1^+ - u_1^- \\
(a_0 - 3a_1 + 9a_2 - 92) = u_2^+ - u_2^- \\
(a_0 - a_1 + a_2 - 96) = u_3^+ - u_3^- \\
(a_0 - 98) = u_4^+ - u_4^- \\
(a_0 + a_1 + a_2 - 100) = u_5^+ - u_5^- \\
u_i^+, u_i^- \geq 0, \text{ for all } i.
\]

One of the optimum solutions of this model is \( \hat{\mathbf{a}} = (\hat{a}_0, \hat{a}_1, \hat{a}_2) = (98.5, 1, -0.5) \), so the fit given by this solution is 
\[
f(t, \hat{\mathbf{a}}) = 98.5 + t - 0.5t^2,
\]
with a residue of 1, in \( L_\infty \)-measure units. The corresponding values of positive and negative parts of the deviations in this optimum solution are \( \hat{u}^+ = (1, 0, 1, 0.5, 1) \) and \( \hat{u}^- = (0, 1, 0, 0, 0) \), and it can be verified that this optimum solution satisfies 
\[
(u_i^+)(u_i^-) = 0 \text{ for all } t.
\]
Even though this Model 1 is a perfectly valid LP model for the problem of minimizing the $L_\infty$-measure of deviation, it has the disadvantage of using the variables $u_i^+, u_i^-$ representing the positive and negative parts of deviations in additional constraints in the model, as explained earlier.

A more direct model for the problem of minimizing $L_\infty(a)$ is the following Model 2 given below. As explained earlier, Model 2 is the better model to use for minimizing $L_\infty(a)$. One of the optimum solutions for this model is the same $\hat{a}$ that was given as the optimum solutions for Model 1, so it leads to the same fit $f(t, \hat{a})$ as described under Model 1.

**Model 2:**

Minimize $z$

subject to

\[-z \leq (a_0 - 5a_1 + 25a_2 - 80) \leq z\]
\[-z \leq (a_0 - 3a_1 + 9a_2 - 92) \leq z\]
\[-z \leq (a_0 - a_1 + a_2 - 96) \leq z\]
\[-z \leq (a_0 - 98) \leq z\]
\[-z \leq (a_0 + a_1 + a_2 - 100) \leq z\]
\[z \geq 0\]

All three methods, the $L_2$, $L_1$, $L_\infty$ methods lead to reasonably good fits for the yield in this chemical reaction, so any one of these fits can be used as the functional form for yield when the reaction temperature is in the range used under this experiment.

2.4.5 Minimizing Positive Combinations of Excesses/Shortages

In many systems, the decision makers usually set up target values for one or more linear functions of the decision variables whose values characterize the way the system operates. Suppose the decision variables are $x = (x_1, \ldots, x_n)^T$ and a linear function $\sum a_j x_j$ has a target value of $b$.

Targets may be set up for many such linear functions. If each of these desired targets is included as a constraint in the model, that model may not have a feasible solution either because there are too many constraints in it, or because some target constraints conflict with the others. That is why in these situations one does not normally require that the target values be met exactly. Instead, each linear function with a target value is allowed to take any value, and a solution that minimizes a penalty function for deviations from the targets is selected for implementation.
For the linear function \( \sum a_jx_j \) with target value \( b \), the **excess** at the solution point \( x \) (or the **positive part of the deviation** \( (\sum a_jx_j - b)^+ \)) and the **shortage** at \( x \) (the **negative part of the deviation** \( (\sum a_jx_j - b)^- \)) are defined to be

- If \( (\sum a_jx_j - b) \geq 0 \), excess \( (\sum a_jx_j - b)^+ = (\sum a_jx_j - b) \).
- Shortage \( (\sum a_jx_j - b)^- = 0 \)

- If \( (\sum a_jx_j - b) \leq 0 \), excess \( (\sum a_jx_j - b)^+ = 0 \).
- Shortage \( (\sum a_jx_j - b)^- = |(\sum a_jx_j - b)| \).

Therefore, both excess and shortage are always \( \geq 0 \), and the penalty term corresponding to this target will be \( \alpha (\sum a_jx_j - b)^+ + \beta (\sum a_jx_j - b)^- \), where \( \alpha, \beta \geq 0 \) are, respectively, the penalties per unit excess, shortage (\( \alpha, \beta \) may not be equal, in fact one of them may be positive and the other 0) set by the decision makers.

The **penalty function** = sum of the penalty terms corresponding to all the targets, by minimizing it subject to the essential constraints on the decision variables, we can expect to get a compromise solution to the problem. If it makes the deviations from some of the targets too large, the corresponding penalty coefficients can be increased and the modified problem solved again. After a few iterations like this, one usually gets a reasonable solution for the problem.

The minimum value of the penalty function is \( \geq 0 \), and it will be 0 iff there is a feasible solution meeting all the targets. When there is no feasible solution meeting all the targets, the deviations from some targets will always be nonzero; minimizing the penalty function in this case seeks a balance among the various deviations from the targets, that is, it seeks a good compromise solution.

By expressing the deviation \( (ax - b) \), which may be positive or negative, as the difference \( u^+ - u^- \) of two nonnegative variables, the excess \( (ax - b)^+ \) defined above will be \( u^+ \) and the shortage \( (ax - b)^- \) defined above will be \( u^- \) as long as the condition \((u^+)(u^-) = 0\) holds. For this, remember the precautions expressed in the Cautionary Note 2.1 given above.

**Example 2.9.** We provide an example in the context of a simple transportation problem. Suppose a company makes a product at two plants \( P_i, i = 1, 2 \). At plant \( P_i \), \( a_i \) (in tons) and \( g_i \) (in $/ton) are the production capacity and production cost during regular time working hours; and \( b_i \) (in tons) and \( h_i \) (in $/ton) are the production capacity and production cost during overtime working hours.

The company has dealers in three markets, \( M_j, j = 1, 2, 3 \) selling the product. The selling price in different markets is different. In market \( M_j \), the estimated demand is \( d_j \) (in tons), and up to this demand of \( d_j \) tons can be sold at the selling price of \( p_j \) (in $/ton), beyond which the market is saturated. However, in each market \( j \), there are wholesalers who are willing to buy any excess over the demand at the price of \( s_j \) (in $/ton).
The cost coefficient $c_{ij}$ (in $/ton$) is the unit transportation cost for shipping the product from plant $i$ to market $j$. All this data is given in the following table.

<table>
<thead>
<tr>
<th>$i$</th>
<th>$j$</th>
<th>$c_{ij}$</th>
<th>$a_i$</th>
<th>$b_i$</th>
<th>$g_i$</th>
<th>$h_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>11</td>
<td>900</td>
<td>300</td>
<td>100</td>
<td>130</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>7</td>
<td>500</td>
<td>200</td>
<td>120</td>
<td>160</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>400</td>
<td>150</td>
<td>140</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>135</td>
<td>137</td>
<td>130</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We want to formulate the problem of finding the best production, shipping plan to maximize net profit ($=$sales revenue − production costs), as an LP. There is no requirement that the amount shipped to any of the markets should equal or exceed the demand at it, in fact any amount of the available product can be shipped to any of the markets. Clearly the decision variables in this problem are $x_{ij} =$ tons shipped from $P_i$ to $M_j$; $i = 1, 2$; $j = 1, 2, 3$ $y_i =$ tons produced in $P_i$, $i = 1, 2$ $y_{i1}, y_{i2} =$ tons of regular, overtime production at $P_i$, $i = 1, 2$.

The essential constraints in this problem are the production capacity constraints, these cannot be violated. They are

$$
x_{11} + x_{12} + x_{13} = y_1 = y_{11} + y_{12}
\quad x_{21} + x_{22} + x_{23} = y_2 = y_{21} + y_{22}
\quad 0 \leq y_{i1} \leq a_i, \quad 0 \leq y_{i2} \leq b_i
\quad \text{for } i = 1, 2
$$

From the production costs, we see that the slope of the production cost function at each plant is monotonic increasing, hence it is PL convex and its negative is PL concave. So, this negative production cost that appears as a term in the overall objective function to be maximized can be expressed as $-(g_{11}y_{11} + h_{11}y_{12} + g_{21}y_{21} + h_{21}y_{22})$.

The demand $d_j$ at market $j$ is like a target value to ship to that market, but the actual amount sent there can be anything. For each unit of excess sent over the demand, there is a drop in the sales revenue of $(p_j - s_j)/unit$. So the total sales revenue can be expressed as $(\sum_{i=1}^{2} x_{ij})p_j - (\sum_{i=1}^{2} x_{ij} - d_j)^+(p_j - s_j)$. So, our problem is

$$
\text{Maximize } \sum_{j=1}^{3} \left[ \left( \sum_{i=1}^{2} x_{ij} \right) p_j - \left( \sum_{i=1}^{2} x_{ij} - d_j \right)^+ (p_j - s_j) \right] - (g_{11}y_{11} + h_{11}y_{12} + g_{21}y_{21} + h_{21}y_{22}) - \sum_{i=1}^{2} \sum_{j=1}^{3} c_{ij}x_{ij}
$$
subject to the constraints (2.14). Putting it in minimization form and linearizing, it is

\[
\text{Minimize } (g_1 y_{11} + h_1 y_{12} + g_2 y_{21} + h_2 y_{22}) + \sum_{i=1}^{2} \sum_{j=1}^{3} c_{ij} x_{ij} \\
- \sum_{j=1}^{3} \left[ \left( \sum_{i=1}^{2} x_{ij} \right) p_j - u_j^+ (p_j - s_j) \right]
\]

Subject to (2.14), and

\[
\sum_{j=1}^{2} x_{ij} = d_j + u_j^+ - u_j^- \quad \text{for all } j \\
u_j^+, u_j^- \geq 0 \quad \text{for all } j.
\]

### 2.5 Multiobjective LP Models

So far we discussed only problems in which there is a single well-defined objective function specified to be optimized. In most real-world decision-making problems there are usually several objective functions to be optimized simultaneously. In many of these problems, the objective functions conflict with one another; that is, moving in a direction that improves the value of one objective function often makes the value of some other objective function worse. See (Charnes and Cooper (1977), Hwang and Masud (1979), Keeney and Raiffa (1976), Sawaragi et al. (1985), Steuer (1986)), for a discussion of multiobjective optimization.

When dealing with such a conflicting set of objective functions, even developing a concept of optimality that every one can agree on has turned out to be very difficult. With the result there is no universally accepted concept of optimality in multiobjective optimization.

Hence, all practical methods for handling multiobjective problems focus on finding some type of a compromise solution.

Let \( x = (x_1, \ldots, x_n)^T \) denote the vector of decision variables. Let \( z_1(x), \ldots, z_k(x) \) denote the \( k \) objective functions to be optimized simultaneously. If any one of them is to be maximized, replace it by its negative, so all the objective functions are to be minimized. Then this multiobjective LP is of the form

Minimize \( z_1(x), \ldots, z_k(x) \) simultaneously

subject to \( Ax = b \quad (2.15) \)

\[
D x \geq d \\
x \geq 0.
\]
It is possible that each objective function is measured in its own special units. A feasible solution \( \bar{x} \) to the problem is said to be a pareto optimal solution (various other names used for the same concept are: vector minimum, nondominated solution, equilibrium solution, efficient solution, etc.) to (2.15) if there exists no other feasible solution \( x \) that is better than \( \bar{x} \) for every objective function and strictly better for at least one objective function; that is, if there exists no feasible solution \( x \) satisfying
\[
\begin{align*}
    z_r(x) & \leq z_r(\bar{x}) \quad \text{for all } r = 1 \text{ to } k \text{; and} \\
    z_r(x) & < z_r(\bar{x}) \quad \text{for at least one } r.
\end{align*}
\]

A feasible solution that is not a nondominated solution is called a dominated solution to the problem. Clearly, a dominated solution is never a desirable solution to implement, because there are other solutions better than it for every objective function. So for a feasible solution to be a candidate to be considered for (2.15), it must be a nondominated solution only.

**Nobel Prize in This Area:** The mathematical theory of nondominated solutions is very highly developed. John Nash was awarded the 1994 Nobel Prize in economics for proving the existence of nondominated solutions for certain types of multiobjective problems, and a highly popular Hollywood movie “A Beautiful Life” has been made based on his life.

Very efficient algorithms have been developed for enumerating the set of all nondominated solutions to multiobjective LPs; this set is commonly known as the efficient frontier. However, typically there are far too many nondominated solutions to multiobjective LPs, and so far no one has been able to develop a concept for the best among them, or an efficient way to select an acceptable one. So, much of the highly developed mathematical theory on nondominated solutions remains unused in practice.

**Example 2.10.** Consider a multiobjective LP in which two objective functions \( z(x) = (z_1(x), z_2(x)) \) are required to be minimized simultaneously. Suppose \( \bar{x} \) with objective values \( z(\bar{x}) = (100, 200) \) and \( \hat{x} \) with \( z(\hat{x}) = (150, 180) \) are two nondominated feasible solutions for this problem. The solution \( \hat{x} \) is a better solution than \( \bar{x} \) for objective function \( z_1(x) \), but \( \bar{x} \) is better than \( \hat{x} \) for \( z_2(x) \). In this pair, improvement in the value of \( z_1(x) \) comes at the expense of deterioration in the value of \( z_2(x) \), and it is not clear which solution is better among these two.

The question can be resolved if we can get some quantitative compromise (or tradeoff) information between the two objectives; that is, how many units of \( z_2(x) \) are the decision makers willing to sacrifice to improve the value of \( z_1(x) \) by one unit? Unfortunately, such compromise information is not available in multiobjective problems; that is what makes them hard to solve.

As another illustration, consider a problem in which two objective functions \( z_1, z_2 \) are required to be minimized simultaneously. If \( \bar{x} \) is a feasible solution to the problem with values \( \bar{z}_1, \bar{z}_2 \) for the two objective functions, we represent \( \bar{x} \) by the point \( (\bar{z}_1, \bar{z}_2) \) in the \( z_1, z_2 \)-plane. In Fig. 2.8, we mark the points in the \( z_1, z_2 \)-plane corresponding to feasible solutions of the problem. They form the dotted region in
Fig. 2.8 Dotted region consists of points in the objective plane corresponding to feasible solutions. The point \( \hat{z} \) does not correspond to a pareto optimum point, since points in the cone region marked by the angle sign, which are in the dotted area, correspond to strictly superior feasible solutions on one or both objective functions. The thick boundary curve corresponds to the efficient frontier.

The reader should not be fooled by the word *optimum* in the phrase *pareto optimum*. In a multiobjective model, a pareto optimum does not have the nice optimality properties that we have seen in single objective models. Remember that a pareto optimum point is just a feasible solution with the property that any move from it, if it leads to a gain in the value of one objective function, it also leads to a loss in the value of another objective function. Usually there are many such points, and it is hard to determine which efficient solution is better unless we have some idea of how much one unit decrease in the value of \( z_2 \) is worth in terms of units of \( z_1 \).

### 2.5.1 Practical Approaches for Handling Multiobjective LPs in Current Use

As pointed out earlier, if complete compromise (or exchange, or tradeoff) information between unit values of the various objective functions is available, it will make it much easier to handle the multiobjective problem. Considering (2.15), suppose the decision makers determine that \( c_1 (=1) \) units of \( z_1(x) \) (in whatever units this objective function is measured in) is equivalent to (or has the same merit or value as)
2.5 Multiobjective LP Models

c_2 units of \( z_2(x) \) (in its own units), that is equivalent to \( c_3 \) units of \( z_3(x) \), \ldots, which is also equivalent to \( c_k \) units of \( z_k(x) \). This vector \( c = (c_1, \ldots, c_k) \) gives complete compromise or exchange information between the various objective functions in this problem, and so can be called the exchange vector. As \( z_r(x), c_r \) are in the same units, the quantity \( (1/c_r)z_r(x) \) is a dimensionless quantity, and we can form the sum \( \sum_{r=1}^{k} (1/c_r)z_r(x) \) and use it as a single objective function that measures the value of the solution vector \( x \). Hence, given the exchange vector \( c \), the multiobjective problem (2.15) is equivalent to the single objective problem of minimizing \( \sum_{r=1}^{k} (1/c_r)z_r(x) \) subject to the constraints in (2.15).

Unfortunately, in many real-world applications, this exchange vector is not available. Usually there may be several decision makers interested in the solution of this multiobjective problem, and each one may have a different opinion of what the value of the exchange coefficient \( c_r \) should be for each \( r \). So, there is no universal agreement on the exchange vector, and the challenge is to obtain a satisfactory solution of the multiobjective problem, without explicitly using any exchange vector.

Even though the practical approaches in use for handling multiobjective problems do not mention exchange vectors directly, they get it indirectly using different wording that the various decision makers find easier to answer.

2.5.2 Weighted Average Technique

This technique uses the tradeoff information in the form of what are called weights measuring the relative importance of the various objective functions, and these weights can be interpreted also as cost coefficients attached to unit values of the various objective functions. The process of generating these weights will be easier if all the objective functions are transformed and measured in common units, say money units, scores, etc.

Let \( w = (w_1, \ldots, w_k) \) be the vector of weights given. From the discussion above, forming the sum \( \sum_{r=1}^{k} w_r z_r(x) \) makes sense, and this technique takes the solution of the multiobjective LP (2.15) to be an optimum solution of the single objective LP:

\[
\text{Minimize } z(x) = \sum_{r=1}^{k} w_r z_r(x) \\
\text{subject to } Ax = b \\
Dx \geq d \\
x \geq 0.
\]  

(2.16)

It can be shown that if all \( w_r > 0 \), then every optimum solution of (2.16) is a nondominated solution for (2.15). So, this type of optimizing a positive weighted combination of all the objective functions is commonly used to generate a nondominated solution for the problem. But the solution obtained depends critically on the choice of the weights \( w_1, \ldots, w_k \) used in combining the original objective functions \( z_1(x), \ldots, z_k(x) \) into the composite objective function \( z(x) \) in (2.16).
There may be several decision makers who have a stake in determining the optimum solution to be selected for implementation. They may not all agree on the choice of the weight vector to be used. It usually takes a lot of planning, discussion, and negotiations, and many compromises, before a weight vector that everyone can agree upon is arrived at. For this negotiation process, it is often helpful to solve (2.16) with a variety of weight vectors and review the optimum solutions that come up, before selecting one of them for implementation.

Example 2.11. Consider the fertilizer problem discussed in Example 2.7, in which the constraints on the decision variables $x_1, x_2$ = tons of hi-ph, lo-ph fertilizer made daily are

$$2x_1 + x_2 \leq 1,500, \quad x_1 + x_2 \leq 1,200$$
$$x_1 \leq 500, \quad x_1, x_2 \geq 0.$$ 

For hi-ph, lo-ph, the selling prices are $222, 107/ton, respectively; and the net profit coefficients are $15, 10, respectively. The important objectives all to be maximized are net profit $z_1(x) = 15x_1 + 10x_2$, total sales revenue (used as a measure of market share); $z_2(x) = 222x_1 + 107x_2$, sales revenue from hi-ph sales (used as a measure of hi-tech market share); $z_3(x) = 222x_1$; all measured in units of Dollar. The multiobjective problem is to maximize $z_1(x), z_2(x), z_3(x)$ simultaneously, subject to the constraints on $x_1, x_2$ given above. Suppose the decision makers have decided that the weights for the objective functions $z_1(x), z_2(x), z_3(x)$ (measuring their relative importance) are 0.5, 0.25, 0.25, respectively. Then we take a compromise solution for this multiobjective problem to be an optimum solution of the single objective function LP:

Maximize $0.5(15x_1 + 10x_2) + 0.25(222x_1 + 107x_2) + 0.25(222x_1)$

subject to $2x_1 + x_2 \leq 1,500, \quad x_1 + x_2 \leq 1,200$ 
$$x_1 \leq 500, \quad x_1, x_2 \geq 0.$$ 

2.5.3 The Goal Programming Approach

The goal programming approach is perhaps the most popular method used for handling multiobjective problems in practice. It has the added conveniences that different objective functions can be measured in different units, and that it is not necessary to have all the objective functions in the same (either maximization or minimization) form. This method developed by A. Charnes has nice features that appeal to the intuition of business people; that is why it is the common method in usage. Several other references on goal programming are given at the end of this chapter (Charnes and Cooper (1977), Hwang and Masud (1979), Keeney and Raiffa (1976), Sawaragi et al. (1985), Schniederjans (1995), Sponk (1981), Steuer 1986).
The most appealing feature of this method is that instead of trying to optimize each objective function, the decision maker is asked to specify a goal or target value that realistically is the most desirable value for that function (the name of the method comes from this feature). Considering (2.15), we will denote the goal selected for \( z_r(x) \) by \( g_r \) for \( r = 1 \) to \( k \). The decision makers are also required to specify a unit penalty coefficient \( \alpha_r \geq 0 \) for each unit the value of \( z_r(x) \) is in excess of the goal \( g_r \), and a unit penalty coefficient \( \beta_r \geq 0 \) for each unit the value of \( z_r(x) \) is short of the goal \( g_r \). These penalty coefficients play the role of exchange or tradeoff coefficients between the various objective functions discussed earlier in this method.

In terms of this goal setting, the objective functions are divided into three types:

**Type 1: Those for which the higher the value the better:** Each of these objective functions should really be maximized; for each of them the goal is like a minimum acceptable value for it. Objective values \( \geq \) the goal are the most desirable; those below the goal are to be avoided as far as possible, and are penalized with positive penalties. So, for objective functions \( z_r(x) \) of this type, \( \alpha_r = 0 \) and \( \beta_r > 0 \).

**Type 2: Those for which the lower the value the better:** These objective functions should be minimized, for them the goal is like a maximum acceptable value. Objective values \( \leq \) goal are desirable, those \( > \) the goal are penalized. So for \( z_r(x) \) of this type, \( \alpha_r > 0 \) and \( \beta_r = 0 \).

**Type 3: Those for which the preferred value is the goal:** For these objective functions their goal is the most desirable value, and both deviations above or below the goal are penalized, So, for objective functions \( z_r(x) \) of this type, both \( \alpha_r > 0 \) and \( \beta_r > 0 \).

At any feasible solution \( x \), for \( r = 1 \) to \( k \), we express the deviation in the \( r \)th objective function from its goal, \( z_r(x) - g_r \), as a difference of two nonnegative variables

\[
z_r(x) - g_r = u_r^+ - u_r^- \quad u_r^+, u_r^- \geq 0,
\]

where \( u_r^+, u_r^- \) are the **positive** and **negative parts of the deviation** \( z_r(x) - g_r \) as explained earlier. That is, \( u_r^+ = \max \{0, z_r(x) - g_r\} \) and \( u_r^- = \max \{0, -(z_r(x) - g_r)\} \).

Given this information, the goal programming approach takes the solution of the multiobjective problem (2.15) to be a feasible solution that minimizes the penalty function \( \sum_{r=1}^{k} (\alpha_r u_r^+ + \beta_r u_r^-) \). So, it takes the solution for (2.15) to be an optimum solution of the single objective LP.

\[
\text{Minimize} \quad \sum_{r=1}^{k} (\alpha_r u_r^+ + \beta_r u_r^-)
\]

subject to

\[
z_r(x) - g_r = u_r^+ - u_r^- \quad r = 1 \text{ to } k \tag{2.17}
\]

\[
Ax = b, \quad Dx \geq d
\]

\[
u_r^+, u_r^- \geq 0, \quad r = 1 \text{ to } k.
\]
As all $\alpha_r$ and $\beta_r \geq 0$, and from the manner in which the values for $\alpha_r, \beta_r$ are selected, an optimum solution of this problem will try to meet the targets set for each objective function or deviate from them in the desired direction as far as possible. If the optimum solution obtained for this problem is not considered satisfactory for (2.15), the search for a better solution can be continued using this same single objective LP model with revised goals, or penalty coefficients, or both.

It can be shown that this goal programming approach is equivalent to the positive linear combination approach when all the objective functions $z_r(x)$ are linear.

**Example 2.12.** Consider the multiobjective problem of the fertilizer manufacturer discussed in Examples 2.7 and 2.11. Suppose the first objective function $z_1(x) = \text{net daily profit} = $(15x_1 + 10x_2)$ with a goal of $13,000 is a Type 1 objective coefficient with penalty coefficients for excess, shortage of $\alpha_1 = 0, \beta_1 = 0.5$.

Suppose the second objective function $z_2(x) = \text{market share}$, now measured by the daily fertilizer tonnage sold $= (x_1 + x_2)$ tons, with a goal $g_2 = 1,150$ tons, is also a Type 1 objective function with penalty coefficients for excess, shortage of $\alpha_2 = 0, \beta_2 = 0.3$.

Suppose the third objective function, $z_3(x) = \text{hi-tech market share}$ now measured by daily hi-ph tonnage sold $= x_1$ tons, is a Type 3 objective function with a goal $g_3 = 400$ tons, with penalty coefficients for excess, shortage of $\alpha_3 = 0.2, \beta_3 = 0.2$.

With this data, the goal programming model for this problem is

$$\begin{align*}
\text{Minimize} & \quad 0.5u_1^- + 0.3u_2^- + 0.2(u_3^- + u_3^+) \\
\text{subject to} & \\
15x_1 + 10x_2 + u_1^- - u_1^+ = 13,000 & \\
x_1 + x_2 + u_2^- - u_2^+ = 1,150 & \\
x_1 + u_3^- - u_3^+ = 400 & \\
2x_1 + x_2 & \leq 1,500 \\
x_1 + x_2 & \leq 1,200 \\
x_1 & \leq 500 \\
x_1, x_2, u_1^-, u_1^+, u_2^-, u_2^+, u_3^-, u_3^+ & \geq 0
\end{align*}$$

An optimum solution of this problem is:

$\hat{x} = (\hat{x}_1, \hat{x}_2)^T = (350, 800)^T$.

The solution $\hat{x}$ attains the goals set for net daily profit, and total fertilizer tonnage sold daily, but falls short of the goal on the hi-ph tonnage sold daily by 50 tons. The vector $\hat{x}$ is the solution for this multiobjective problem obtained by goal programming, with the goals and penalty coefficients given above.
2.6 Exercises

2.1. Hiring workers for a new project: A contractor is working on a project expected to last an year divided into six periods. He needs workers for this project, who can be hired in two different ways:

**Steady workers:** These people are hired at the beginning of the year for the whole year, they get paid at the rate of $c_1$/period for each of the six periods of the year ($c_1 = 20,000$)

**Casual workers:** These people are hired at the beginning of each period for that period only. As the availability of people seeking employment in the various periods varies, the rate of pay per casual worker in period $t$ is expected to be $c_2_t$ in period $t$ for $t = 1–6$ ($c_{2_t} = 13,000, 17,000, 23,000, 25,000, 21,000, 15,000$, respectively, for $t = 1–6$).

The company estimates that they will need $d_t$ workers in period $t$ for $t = 1–6$ ($d_t = 15, 20, 23, 17, 25, 28$, respectively, for $t = 1–6$).

In any period $t$, if the number of workers on the payroll in that period exceeds $d_t$, then the excess workers will have to be paid their contracted salary, but they will essentially be idle in that period. To enforce responsible hiring, the Head Office charges the contractor an amount of $2,500 penalty/idle worker/period.

Formulate the problem of determining the optimal hiring policy to meet worker requirements to minimize the total cost + penalty, as a linear program, ignoring the integer restrictions on the number of workers hired.

2.2. Blend for foundry sand: A steel foundry uses a lot of sand for preparing molds for castings. They have five different suppliers for sand denoted by $S_1$ to $S_5$, each has an upper limit that they can supply per week. Sand is classified into four different particle sizes; the supply from each supplier has a different particle size composition. For the castings the company produces, the ideal particle size composition of sand is known, and the demand for sand at the company is 800 tons/week.

Data for the following items is tabulated below.

- $f_{ij}$ = fraction by weight of particle size $j$ in sand supplied by $S_i$, $i = 1–5$, $j = 1$ to 4
- $U_i$ = upper limit on tons of sand that $S_i$ can supply per week, $i = 1–5$
- $g_j$ = fraction by weight of particle size $j$ in the ideal sand for the company, $j = 1–4$

This data is tabulated below.

<table>
<thead>
<tr>
<th>Supplier $i$</th>
<th>$f_{ij}$ For $j = $</th>
<th>$U_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>$S_1$</td>
<td>0.10</td>
<td>0.40</td>
</tr>
<tr>
<td>$S_2$</td>
<td>0.20</td>
<td>0.10</td>
</tr>
<tr>
<td>$S_3$</td>
<td>0.30</td>
<td>0.58</td>
</tr>
<tr>
<td>$S_4$</td>
<td>0.40</td>
<td>0.15</td>
</tr>
<tr>
<td>$S_5$</td>
<td>0.60</td>
<td>0.17</td>
</tr>
</tbody>
</table>

$g_j$ = 0.45 0.30 0.15 0.10 demand = 800
Clearly each supplier’s sand deviates from the desired ideal sand. So, the company has decided to blend the sands from the various suppliers into a mixture and use it.

Write down the decision variables needed to model the problem of finding an optimal blend that matches the ideal particle size composition as closely as possible. Develop a measure of deviation of the prepared mixture’s particle size composition from the ideal, and formulate the problem of minimizing this measure of deviation as a linear program to find the best blend.

2.3. Investment fund allocation among different shopping malls: A company wants to invest their funds among different shopping malls in the Atlanta retailing system. Their ultimate goal is of course to maximize the expected return from their investment. They believe that they can achieve this by making sure that the sales volumes in the shopping malls in which they invest remain high. Their research has shown that a major factor affecting the sales volume in a shopping mall is its patronization rate, which is defined as the share of the shopping trips in the area that it attracts, and this can be estimated through observations.

The company has investigated 15 shopping malls in the Atlanta area. Of these, they have decided to consider only shopping malls that attract 7% or more of the total shopping trips in the Atlanta area for investment. That narrowed the list to only five shopping malls.

Two other major factors affecting the sales volume at a shopping mall are the average income level (annual income per member of household) of families in the neighborhood of the mall and the number of major tenants in the mall. The following table gives the data on these major factors for the five selected malls. In this table, PPR$_j$ = predicted patronage rate of mall $j$, as a percentage of the total of the five malls; AI$_j$ = average annual income level per household member in the neighborhood of mall $j$ in Dollar; MT$_j$ = number of major tenants in mall $j$, for $j = 1–5$.

<table>
<thead>
<tr>
<th>Mall $j$</th>
<th>PPR$_j$</th>
<th>AI$_j$</th>
<th>MT$_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cumberland</td>
<td>20.5</td>
<td>19,700</td>
<td>4</td>
</tr>
<tr>
<td>2. Greenbriar</td>
<td>16.3</td>
<td>10,500</td>
<td>2</td>
</tr>
<tr>
<td>3. Lenox Square</td>
<td>25.3</td>
<td>18,800</td>
<td>3</td>
</tr>
<tr>
<td>4. Northlake</td>
<td>22.2</td>
<td>16,700</td>
<td>3</td>
</tr>
<tr>
<td>5. West End</td>
<td>15.7</td>
<td>8,200</td>
<td>1</td>
</tr>
</tbody>
</table>

The decision variables to be determined are the fractions $x_j$ of their total investment to allocate to mall $j$, for $j = 1–5$. They want to determine these decision variables to meet the following three goals as closely as possible (all these goals have the same level of importance).

**Goal 1:** The $x_j$’s should be proportional to the PPR$_j$’s as closely as possible.

**Goal 2:** The $x_j$’s should be proportional to the AI$_j$’s as closely as possible.

**Goal 3:** The $x_j$’s should be proportional to the MT$_j$’s as closely as possible.
2.6 Exercises

Formulate the problem of determining the decision variables to meet all these goals as closely as possible, using goal programming (Khorrampishshgol and Okoruwa 1994).

2.4. Finding an optimum mix of water from different underground wells for people in the Gaza Strip: WHO has set up standards of 250 mg \( \ell^{-1} \) for chlorides and 50 mg \( \ell^{-1} \) for nitrates as upper bounds for these chemicals in drinking water. The Gaza Strip is a narrow strip along the Mediterranean sea, with a population of around 1.1 million people, which has serious water quality problems not only for agricultural and industrial use, but also for drinking. Groundwater is the only significant source of water they have. Because of waste water contamination, uncontrolled use of agricultural fertilizer and pesticides and industrial pollutants, which directly penetrate to the groundwater reservoir through pores in the rocks, and indirectly by decomposition, salts in some wells are at high levels. For example, level of nitrates in the water in some wells exceeds 300 mg \( \ell^{-1} \) (six times WHO upper limit), and the level of chlorides may exceed WHO limit by four to five times. Drinking water with nitrate level over 150 mg \( \ell^{-1} \) poses an extreme risk of blue baby syndrome in infants and carcinogenic effects in adults. Similarly, high levels of chlorides in water make it unacceptable for drinking due to the salinity it causes, and also causes high blood pressure in those who drink it.

<table>
<thead>
<tr>
<th>No.</th>
<th>Well name</th>
<th>Capacity ( Q ) (m(^3) h(^{-1}))</th>
<th>Chlorides mg ( \ell^{-1} )</th>
<th>Nitrates mg ( \ell^{-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sheikh Radwan 1</td>
<td>180</td>
<td>273</td>
<td>135</td>
</tr>
<tr>
<td>2</td>
<td>Sheikh Radwan 1A</td>
<td>180</td>
<td>245</td>
<td>140</td>
</tr>
<tr>
<td>3</td>
<td>Sheikh Radwan 3</td>
<td>150</td>
<td>1,015</td>
<td>135</td>
</tr>
<tr>
<td>4</td>
<td>Sheikh Radwan 4</td>
<td>180</td>
<td>1,085</td>
<td>90</td>
</tr>
<tr>
<td>5</td>
<td>Sheikh Radwan 7</td>
<td>180</td>
<td>553</td>
<td>175</td>
</tr>
<tr>
<td>6</td>
<td>Sheikh Radwan 7A</td>
<td>180</td>
<td>485</td>
<td>115</td>
</tr>
<tr>
<td>7</td>
<td>Sheikh Radwan 8</td>
<td>150</td>
<td>133</td>
<td>80</td>
</tr>
<tr>
<td>8</td>
<td>Sheikh Radwan 9</td>
<td>190</td>
<td>133</td>
<td>90</td>
</tr>
<tr>
<td>9</td>
<td>Sheikh Radwan 10</td>
<td>190</td>
<td>90</td>
<td>60</td>
</tr>
<tr>
<td>10</td>
<td>Sheikh Radwan 11</td>
<td>190</td>
<td>110</td>
<td>80</td>
</tr>
<tr>
<td>11</td>
<td>Sheikh Radwan 12</td>
<td>180</td>
<td>110</td>
<td>45</td>
</tr>
<tr>
<td>12</td>
<td>Sheikh Radwan 13</td>
<td>180</td>
<td>460</td>
<td>200</td>
</tr>
<tr>
<td>13</td>
<td>Sheikh Radwan 15</td>
<td>180</td>
<td>110</td>
<td>45</td>
</tr>
<tr>
<td>14</td>
<td>Sheikh Radwan 16</td>
<td>180</td>
<td>110</td>
<td>45</td>
</tr>
<tr>
<td>15</td>
<td>Sheikh Ejleen 1</td>
<td>150</td>
<td>810</td>
<td>110</td>
</tr>
<tr>
<td>16</td>
<td>Sheikh Ejleen 2</td>
<td>120</td>
<td>440</td>
<td>80</td>
</tr>
<tr>
<td>17</td>
<td>Saffa well 1</td>
<td>200</td>
<td>600</td>
<td>225</td>
</tr>
<tr>
<td>18</td>
<td>Saffa well 2</td>
<td>150</td>
<td>400</td>
<td>225</td>
</tr>
<tr>
<td>19</td>
<td>Saffa well 3</td>
<td>100</td>
<td>740</td>
<td>215</td>
</tr>
<tr>
<td>20</td>
<td>Saffa well 4</td>
<td>180</td>
<td>590</td>
<td>110</td>
</tr>
</tbody>
</table>
This question deals with achieving an optimum strategy for mixing the water from different wells to produce drinking water for the people in Gaza, which meets WHO limits for nitrates and chlorides. There are 20 municipal wells as sources for the water; data on the capacity of each well, nitrate, chloride level in the water from each well are given above.

Develop a mathematical model to determine how much water from each of the wells should be drawn (subject to the capacity constraint) and mixed to produce drinking water at the rate of 450 m$^3$ h$^{-1}$, satisfying the WHO limits on nitrates and chlorides as closely as possible (Agha 2006).

2.5. Locations for new water reservoirs: A small country has a more or less rectangular shape, suppose it is represented by a rectangle in the nonnegative orthant of the $x$, $y$-Cartesian plane with $0 \leq x \leq 10$ and $0 \leq y \leq 40$.

The country’s water supply comes from six underground wells located at points (3.21, 39.41), (3.94, 38.70), (5.09, 34.43), (6.05, 34.50), (6.33, 35.82), (6.64, 36.48), respectively.

The country has decided to set up three different holding reservoirs of equal capacity to hold the water pumped from the wells. The total quantity pumped into each of the reservoirs to be set up will be $1.4 \times 10^6$ m$^3$/year. The amounts pumped from the six wells will be 23,546, 311,112, 317,543, 1,182,600, 1,182,600, 1,182,600 m$^3$ per year, respectively.

There are three population centers in the country, each has a municipal water reservoir for supplying water to people in that center. These municipal water reservoirs are located at points (2.23, 10.11), (4.50, 20.22), (7.96, 30.21), respectively.

Because of the topography of the country and the way water distribution pipes are laid, distance between any two points $a = (a_1, a_2)$ and $b = (b_1, b_2)$ for water distribution purposes can be measured using the rectilinear distance between $a$ and $b$, which is $|a_1 - b_1| + |a_2 - b_2|$.

Formulate the problem of finding the locations $(x_i, y_i)$, $i = 1$–$3$ of the three proposed reservoirs to minimize the sum of the rectilinear distances of each of the proposed reservoirs to each of the supply wells and the municipal reservoirs, as an LP.

2.6. How to manage a prime raw material supply curtailment: Coastal States Chemicals and Fertilizers (CSCF) located between Baton Rouge and New Orleans in Louisiana makes eight different products using NG (natural gas) as a prime raw material. Following table provides relevant information on the production of these products at CSCF.
### 2.6 Exercises

<table>
<thead>
<tr>
<th>Product</th>
<th>Prod. capacity (tons day$^{-1}$)</th>
<th>Rate$^a$ (10$^3$FT$^3$ ton$^{-1}$)</th>
<th>NG input ($\text{ton day}^{-1}$)</th>
<th>Net profit ($\text{ton}^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phosphoric acid</td>
<td>400</td>
<td>80</td>
<td>5.5</td>
<td>60</td>
</tr>
<tr>
<td>Urea</td>
<td>250</td>
<td>80</td>
<td>7.0</td>
<td>80</td>
</tr>
<tr>
<td>Ammonium phosphate</td>
<td>300</td>
<td>90</td>
<td>8.0</td>
<td>90</td>
</tr>
<tr>
<td>Ammonium nitrate</td>
<td>300</td>
<td>100</td>
<td>10.0</td>
<td>100</td>
</tr>
<tr>
<td>Chlorine</td>
<td>800</td>
<td>60</td>
<td>15.0</td>
<td>50</td>
</tr>
<tr>
<td>Caustic soda</td>
<td>1,000</td>
<td>60</td>
<td>16.0</td>
<td>50</td>
</tr>
<tr>
<td>Vinyl chloride manomer</td>
<td>500</td>
<td>60</td>
<td>12.0</td>
<td>65</td>
</tr>
<tr>
<td>Hydrofluoric acid</td>
<td>400</td>
<td>80</td>
<td>11.0</td>
<td>70</td>
</tr>
</tbody>
</table>

$^a$Present production rate as a % of capacity.

Cajun Pipeline Co., the main supplier of NG, informed CSCF that they may have to curtail their supply of NG due to shortages in the availability of NG.

It is required to develop contingency plans to determine the new production rates for the various products if there is a (a) 20% and (b) 40% NG supply curtailment. Model the problem for developing these plans optimally as an LP model. Then discuss how to find the impact that natural gas shortages have on company profits (Iverstine and Kinard 1977).

### 2.7. Financial management at NFP (not-for-profit) organizations:

NFPs are usually funded through grants from many different sources, with each individual funding source putting specific restrictions on how and when their grant money can be spent. For example, certain funding sources dictate what maximum and/or minimum percent of a specific employee’s salary can be charged to their grant depending on the work the employee does for the NFP agency. Some sources impose a requirement that funds used from their grant be matched with funds from grants from other specified types of sources.

Here we consider the problem of assigning employee costs and the other costs of the organization to the available grants subject to all the constraints. These costs are classified into three categories: direct, common, and indirect costs.

Direct costs are those that can be traced specifically to a particular program of the NFP (such as the salary of a nurse practitioner working full time in a community clinic run by the NFP. The salary of a medical assistant who works in two departments is also a direct cost split across these two departments according to the proportion of time devoted to each).

Common costs are those nonadministrative costs that are shared across the agency and not easily traced to a particular program; examples are rent, utilities, office supplies, etc.
Indirect costs are management and general administration expenses, like employee benefits, agency management related expenses, etc.

An employee with an FTE (full-time-equivalent) of 1 (i.e., a full-time employee) is either dedicated to one specific program or splits his/her work hours across multiple programs. Managerial employees typically spend a portion of their work hours on specific programs (referred to as program-related FTE) and reminder on agency management (management-related FTE), which will be treated as indirect cost to the agency.

Consider an NFP with four employees working on two programs according to the following data.

<table>
<thead>
<tr>
<th>Name</th>
<th>On prog. FTE</th>
<th>Related to Prog. FTE</th>
<th>Salary $/month</th>
<th>Benefits $/month</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>1</td>
<td>0</td>
<td>5,000</td>
<td>450</td>
</tr>
<tr>
<td>Jill</td>
<td>0.30</td>
<td>0.45</td>
<td>0</td>
<td>300</td>
</tr>
<tr>
<td>Dick</td>
<td>0</td>
<td>0</td>
<td>9,000</td>
<td>900</td>
</tr>
<tr>
<td>Sue</td>
<td>0</td>
<td>0.50</td>
<td>6,000</td>
<td>550</td>
</tr>
<tr>
<td>Total</td>
<td>1.3</td>
<td>0.95</td>
<td>2.25</td>
<td>1.50</td>
</tr>
</tbody>
</table>

The total monthly common costs for the agency are $5,000.

Of the total monthly FTE of 3.75, management-related FTEs are 1.5, or 40% of overall FTEs are dedicated to management. So, 40% of this agency's monthly common costs of $5,000, that is, $2,000, are management related and treated as agency’s indirect costs. The remaining 60% of monthly common costs, that is, $3,000, are program related and treated under direct costs of the agency.

The total monthly indirect costs for this agency comprise of all benefits, managerial salaries, management related common costs, that is, $2,200 + 9,000 + 3,000 + 2,000 = $16,200. The total monthly direct costs for this agency consist of all program related employee salaries and program related common costs = $5,000 + 4,000 + 3,000 + 3,000 = $15,000.

For the next month, this agency has funding from three separate grants. The amounts of these grants and restrictions on how these amounts can be spent are summarized below:

**For Grant 1:** Amount available for the month is $14,000. Restrictions are (1) between 20% and 40% of grant money should be devoted to Program 1 direct costs, (2) between 30% and 60% of grant money should be devoted to Program 2 direct costs.

**For Grant 2:** Amount available for the month is $7,000. Restriction is (1) No more than 40% of grant money to be spent for indirect costs.

**For Grant 3:** Amount available for the month is $10,000. Restriction is (1) At least 40% of grant money to be spent for direct costs.

**For all the grants:** Restriction is: Some overspending for the month above the available amount is allowed, but as far as possible the overspending amount should be kept within 2% of the available amount.
2.6 Exercises

It is required to determine what portion of each employee’s salary and benefits, and what portion of the common costs will be charged to each of the grants for the coming month. Formulate the problem of determining these things as a goal programming problem to satisfy each of the restrictions on spending mentioned above as closely as possible (Mehrotra et al. 2006).

2.8. Advertising media selection problem: There is a lot of interest in media selection for effective advertising. In this problem we consider selection from nine promotional media well known for their objective specific effects. We consider four different objectives, each measured in terms of the number of consumers induced to move from a lower hierarchical stage to a higher level of cognitive development by an advertising event. The four objectives are the following:

1. \( z_1 = \text{Awareness} \): This objective measures the number of new consumers who become aware of the manufacturer’s product or brand as a result of a promotional effort. This objective is most likely to be influenced by network radio and TV jingle slogan campaigns.

2. \( z_2 = \text{Knowledge} \): This measures the number of consumers who acquire product-or-brand related knowledge as a result of the promotional event. This objective is most likely influenced by noncompetitive descriptive copy promoting the merits of the advertised product.

3. \( z_3 = \text{Preference} \): This measures the number of consumers who acquire a preference for the product or brand of interest. This criterion is influenced by competitive copy favorably comparing the advertiser’s brand to competitor’s brands.

4. \( z_4 = \text{Purchase} \): This measures the number of consumers who actually purchase the product sold. Best influenced by retail store copy in the local media.

The various promotional media considered are the following:

1. Slogan/jingle campaign on network radio
2. Slogan/jingle campaign on network TV
3. Descriptive copy in national magazine
4. Descriptive copy on network TV
5. Competitive copy in national magazine
6. Competitive copy on network TV
7. Retail store copy on spot radio
8. Retail store copy on spot TV
9. Retail store copy in top 100 newspapers

Let \( x_j \) denote the number of campaigns to be held on media \( j \), for \( j = 1 \) to \( 9 \); these are the decision variables whose values need to be determined.

Even though all \( x_j \) are integer variables, we will treat them as continuous variables in this problem; for implementation, the noninteger values in the solution will be rounded to nearest integer values.

Each of the above objective function values depend on the decision vector \( x = (x_j) \). For \( i = 1 \) to \( 4 \), \( z_i \) can be estimated by the linear function \( 100(\sum_{j=1}^{9} c_{ij} x_j) \), where data on the \( c_{ij} \) coefficients, desired upper bounds \( u_j \) on the variables \( x_j \), and
cost in Dollar per single campaign of promotional media \( j \), and the total budget are given below (following table gives money figures in $100 units). The \( c_{ij} \) values are estimated through data obtained from marketing surveys.

| \( c_{ij} \) for \( i = 1 \) | Promotional media \( j \) | \( 1 \) | \( 2 \) | \( 3 \) | \( 4 \) | \( 5 \) | \( 6 \) | \( 7 \) | \( 8 \) | \( 9 \) |
|---|---|---|---|---|---|---|---|---|---|
| \( i = 1 \) | \( 150 \) | \( 500 \) | \( 100 \) | \( 100 \) | \( 100 \) | \( 50 \) | \( 50 \) | \( 50 \) | \( c_{ij} \) for Promotional media \( j \) |
| \( i = 2 \) | \( 10 \) | \( 50 \) | \( 200 \) | \( 350 \) | \( 150 \) | \( 10 \) | \( 10 \) | \( 10 \) | \( u_j \) |
| \( i = 3 \) | \( 0 \) | \( 0 \) | \( 150 \) | \( 250 \) | \( 200 \) | \( 350 \) | \( 10 \) | \( 10 \) | \( 10 \) | \( d_j \) |
| \( i = 4 \) | \( 0 \) | \( 10 \) | \( 15 \) | \( 15 \) | \( 20 \) | \( 50 \) | \( 100 \) | \( 200 \) | \( 300 \) |
| \( a_i \) | \( 26 \) | \( 26 \) | \( 12 \) | \( 12 \) | \( 12 \) | \( 12 \) | \( 6 \) | \( 6 \) | \( 6 \) |
| \( b_j \) | \( 69 \) | \( 757 \) | \( 780 \) | \( 1,515 \) | \( 1,100 \) | \( 1,515 \) | \( 100 \) | \( 950 \) | \( 3,500 \) |
| Budget = 35,000. All money figures in $100 units. |

Notice that each of the four objective functions are measured in the same units (number of consumers), so combining them into some type of a composite objective function is perfectly valid.

The mediaplanner of the company wants to use the advertizing budget available most effectively. Formulate this as a multiobjective optimization problem, and discuss a good approach for solving it (Steuer and Oliver 1976).

### 2.9. Time transportation problems with multiple objectives:

Bharat Coking Coal in the state of Bihar, India, gets coal from six different mines, and uses it to make coking coals at six coking plants, each coking plant making a different grade of coking coal.

For \( i = 1 \) to \( 6 \): \( p_i \) denotes the percent of sulfur in the coal from mine \( i \), \( i = 1 \) to \( 6 \); and \( a_i \) is the maximum units of coal that mine \( i \) can ship per period. For \( j = 1 \) to \( 6 \): \( U_j \) denotes the maximum allowable percent of sulfur in the input coal at coking plant \( j \) for \( j = 1 \) to \( 6 \), and \( b_j \) denotes the units of coal required at plant \( j \) per period. For \( i, j = 1 \) to \( 6 \), let \( x_{ij} \) denote the units of coal shipped from mine \( i \) to plant \( j \) in a period, and \( t_{ij} \) the units of time that coal shipped from mine \( i \) takes to reach plant \( j \). This data is given below.

<table>
<thead>
<tr>
<th>Mine ( i ) =</th>
<th>( t_{ij} ) data for plant ( j ) =</th>
<th>( a_i )</th>
<th>( p_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 1 )</td>
<td>( 4 )</td>
<td>( 4 )</td>
<td>( 3 )</td>
</tr>
<tr>
<td>( 2 )</td>
<td>( 8 )</td>
<td>( 2 )</td>
<td>( 4 )</td>
</tr>
<tr>
<td>( 3 )</td>
<td>( 1 )</td>
<td>( 3 )</td>
<td>( 8 )</td>
</tr>
<tr>
<td>( 4 )</td>
<td>( 8 )</td>
<td>( 3 )</td>
<td>( 4 )</td>
</tr>
<tr>
<td>( 5 )</td>
<td>( 4 )</td>
<td>( 2 )</td>
<td>( 4 )</td>
</tr>
<tr>
<td>( 6 )</td>
<td>( 5 )</td>
<td>( 5 )</td>
<td>( 3 )</td>
</tr>
<tr>
<td>( b_j )</td>
<td>( 7 )</td>
<td>( 10 )</td>
<td>( 9 )</td>
</tr>
<tr>
<td>( U_j )</td>
<td>( 0.7 )</td>
<td>( 0.5 )</td>
<td>( 0.7 )</td>
</tr>
</tbody>
</table>
(1) The *time minimizing transportation problem* is the problem of minimizing the time $t_1$, by which shipments from all the mines reach their respective destinations, assuming that all the shipments leave each mine at time point 0 (this type of objective function is encountered commonly in transportation of perishable goods, delivery of emergency supplies, etc.). A single objective transportation problem with this type of objective function is known in the literature as the *bottleneck transportation model*.

In this numerical example we have a bottleneck transportation problem with the additional constraint that the sulfur content of coal delivered at plant $j$ has to be $\leq U_j$ for $j = 1$ to 6.

(a) Formulate this problem.

(b) Notice that the objective function in this model is not linear. Discuss an efficient approach for solving this model (Hint: This can be modeled as a bounded variable LP with variable upper bounds on the variables.).

(2) Now consider the same coal shipping example with multiple time objective functions to be minimized simultaneously, instead of the single time objective function considered earlier. For instance, in this example, let $t_{ij}^3$ be the units of time that a truck going from mine $i$ to plant $j$ takes to reach the highway after leaving mine $i$ (after leaving mine $i$, the truck goes through the congested neighborhood of mine $i$ for some distance before reaching the highway taking it to plant $j$. The time taken for this portion of travel called “congestion time” is $t_{ij}^3$). Also, let $t_{ij}^2$ denote the time that the truck going from mine $i$ to plant $j$ takes for travel between two checkposts on the way. The multiobjective version consider the three time objective functions; those given by the cost data matrices ($t_{ij}^1$), ($t_{ij}^2$), ($t_{ij}^3$) measuring the total travel time, the travel time for travel between two checkposts on the way on each route; and the third measuring the congestion time that each shipment takes to begin highway transit to its destination after leaving the mine where it originates. So, the three objective functions to be minimized are $t_1 = \max \{t_{ij}^1 : (i, j) \text{ such that } x_{ij} > 0\}$, $t_2 = \max \{t_{ij}^2 : (i, j) \text{ such that } x_{ij} > 0\}$, and $t_3 = \max \{t_{ij}^3 : (i, j) \text{ such that } x_{ij} > 0\}$.

Discuss an approach how this multiobjective bottleneck transportation model can be solved (Singh and Saxena 2003).

2.10. Selecting investment opportunities: There are five projects in which money can be invested, covering a 4-year horizon. Details of these projects are given below. In this description, payout refers to either the return on investment or the principal given back.

Project A accepts an investment between $10,000 and $500,000 at the beginning of year 1. This investment is for a 2-year period. The project pays out 30cents/$ invested at the end of the 1st year and $1/$ invested at the end of the second.

Project B is identical to Project A, except that it is available at the beginning of year 2.
Project C is a first year investment available at the beginning of year 1, which pays out $1.10/$ invested at the beginning of year 2. It accepts investments of $20,000 or over.

Project D is a 3-year investment available at the beginning of year 1, which pays out $1.75/$ invested at the beginning of year 4.

Project E becomes available at the beginning of year 3 and will pay out $1.20/$ invested at the beginning of year 4. It accepts an investment of at most $750,000.

Payout received from any of these projects in years 2 and 3 may be reinvested in others, which are available for investment at that time. In addition, short-term (1 year) bank accounts yielding 6% interest are available for any money not invested in the projects in the given year.

Suppose we start year 1 with $1 million of our money to put into a mix of these opportunities, but no more there after, although we will reinvest payouts. The planning horizon considered therefore is beginning of year 1 to beginning of year 4. All cash received at the beginning of year 4 will be withdrawn.

(1) Develop a model to find an investment plan over the planning horizon, which will maximize the total money from the investments by the beginning of year 4
(2) Develop a model to find an investment that maximizes the NPV (net present value) of all the payout money, assuming a discount rate of 10%/year, and that payouts received are not reinvested but cashed out.


2.11. Production planning at a rubber company: Rubicon is a rubber company that used to manufacture a variety of rubber products including tires for forklift trucks and small tractors in the 1960s. At this time the company finds it advantageous to take short lead-time contracts to make small runs of regular automobile snow tires for a large distributor of auto replacement tires. These tires, bearing the distributor’s trademark and made to the distributor’s specifications, utilize surplus capacity at Rubicon. The contract with the distributor calls for a staged delivery schedule of the two types of snow tires (nylon, fiberglass) over the three summer months as indicated below.

<table>
<thead>
<tr>
<th>Delivery schedule</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
</tr>
<tr>
<td>---------------</td>
</tr>
<tr>
<td>31 June</td>
</tr>
<tr>
<td>30 July</td>
</tr>
<tr>
<td>31 August</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Price/tire</td>
</tr>
</tbody>
</table>

Only two types of machines, the Wheeling and Regal machines, can be used in molding tires of the sort covered by the contract. These machines are fully booked until the first of June. After that date, unused capacity would be available spasmodically between other contracts as given below.
The two types of molding machines are similar except that the Wheeling machines are somewhat faster for making both types of tires than the older Regal machines. Here is the data (costs are for the 1960s).

<table>
<thead>
<tr>
<th>Machine</th>
<th>Prod. hr/tire for type</th>
<th>Mc. running cost/hr</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nylon</td>
<td>Fiberglass</td>
</tr>
<tr>
<td>Wheeling</td>
<td>0.15</td>
<td>0.12</td>
</tr>
<tr>
<td>Regal</td>
<td>0.16</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Material costs/tire | $3.10 | $3.90 |
FPS costs/tire | $0.23 | $0.23 |

There is a small storage area adjacent to the production shop where up to one month’s production could be kept until they are delivered to the distributor or to a warehouse where tires can be stored from one month to the next at a cost of $0.10/tire/month. Shipping is scheduled three days prior to the end of the month for delivery before the last day of the month.

(1) Formulate the problem of minimizing the cost of producing and storing the tires to meet this contract on time. Find an optimum solution of your model using any of the available LP software systems.

(2) An additional Wheeling machine was due to arrive at the end of August. For a $200 fee, its arrival can be expedited to a month earlier. This early arrival would make available 172 additional hours of Wheeling machine time in August for this snow tire work.

Discuss how the model for the problem will change if this expediting is carried out, solve the new model, and decide whether the new machine’s arrival should be expedited.

(3) From the optimum solutions, determine a tentative schedule for the maintenance department, indicating when the yearly maintenance check on the various machines could be performed.

(4) The snow tire distributor has found that sales of fiberglass tires had been very good the previous year, so they indicated that they may ask Rubicon to increase their supply of this type of tires. Explain the strategy that Rubicon should adopt if this request comes through (Hint: This needs knowledge of sensitivity analysis discussed later in Chap. 6).

(Source: From Vatter et al. 1978).

2.12. Manpower planning at a mutual life insurance company: At a MLIC (Mutual Life Insurance Company), life insurance is sold through full-time agents
of the company. This exercise refers to the manpower planning effort carried out in the company in 1962 for the 1963–1975 horizon. The most critical element in long-range planning is identifying the company’s personnel requirements, as this is the variable most likely to affect the sales of new life insurance.

A newly hired agent has relatively low productivity (in terms of new life insurance volume generated per year) in the beginning, but their productivity improves with their years of service at the company. Also, there is considerable turnover among the agents, many leave the company after working for some years for other jobs (e.g., out of 100 agents hired in a year, on average only 22 remain with the company 4 years later).

From past records at the company, agent survival rates and agent productivity rates have been estimated as a function of length of service, and this information in terms of \( P_r, S_r, T_r \) is given below, where for \( r = 1, \ldots \)

\[ P_r = \text{The probability that an agent who is in the } r\text{th year of service at the company surviving this year (i.e., continuing in service to go on to the } (r + 1)\text{th year of service)} \]

\[ S_r = \text{Average production rate (i.e., sales of new life insurance commissioned in that year in units of$1,000) of a person in the } r\text{th year of service at the company and surviving into the next year of service} \]

\[ T_r = \text{Average production rate (i.e., sales of new life insurance commissioned in that year in units of$1,000) of a person in the } r\text{th year of service at the company and terminating his service at the company by the end of that year.} \]

<table>
<thead>
<tr>
<th>( r ) = year of service</th>
<th>Money unit = $1,000</th>
<th>( P_r )</th>
<th>( S_r )</th>
<th>( T_r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>0.74</td>
<td>126</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>2nd</td>
<td>0.55</td>
<td>336</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>3rd</td>
<td>0.71</td>
<td>366</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>4th</td>
<td>0.77</td>
<td>404</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>5th</td>
<td>0.79</td>
<td>417</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>6th or higher</td>
<td>0.91</td>
<td>474</td>
<td>116</td>
<td></td>
</tr>
</tbody>
</table>

For \( r \geq 6 \), \( P_r, S_r, T_r \) remain the same.

Using this information, answer the following questions.

1. For \( r = 1 \) to 6 and higher, what is the probability that a person hired at the beginning of first year will survive at the company \( r \) years and goes on into the \((r + 1)\)th year of service?
2. For \( r = 1 \) to 6 and higher, what is the probability that a person hired at the beginning of first year will survive at the company \( r \) years and terminate his service at the company during the \( r \)th year of service?
3. For a person hired at the beginning of first year, define for \( r = 1, \ldots, PC_r \) = the expected sales volume (in units of$1,000) that this person can be expected to generate during the \( r \)th year. Find this \( PC_r \) for \( r = 1, \ldots \).
4. The company has set up sales goals for each of the years 1970–1975. Let these be denoted by \( G_i \) for \( i = 1970–1975 \). Goals for sales volume in several consecutive years in future are set in order to ensure some continuity in sales growth. The company requires that the actual estimated sales in those years must be greater than or equal to these goals.
During the 1962 calendar year, the company hired 430 new agents. For \( i = 1963–1975 \), let \( x_i \) = the number of new agents to be hired during the calendar year \( i \). The company’s objective is to be a growth company, but at the same time does not want to grow so fast as to become unmanageable. To realize these, they want that for each \( i \), \( x_i \) should be between \( (1.5)x_{i-1} \) to \( (1.5)x_{i-1} \).

Ignoring the integer restrictions on the variables \( x_i \), set up a linear programming model to determine \( x_i \) (for \( i = 1963–1975 \)) that minimizes the total number of new agents hired between 1963 and 1975 subject to all the constraints mentioned above.

For \((G_i : i = 1970–1975) = (351, 416, 482, 548, 615, 682)\) and \((636, 712, 793, 874, 958, 1,042)\), respectively, solve the model using some LP software, and compare the two optimum solutions.

(Source: Davis and Webster 1968).

### 2.13. Forecasting expected demand for paint:
Past data of a paint manufacturing company that distributes and sells its paints nationwide, over a 15-year period data on \( \xi \) = home improvement loans granted in the year in the nation as a whole in units of $1 billion, \( \eta \) = population of the country in millions, \( \xi \) = index of building construction started in the nation as a whole, are given above. Formulate the problem of determining \( y \) = its total yearly sales (in $1 million units), as a function of some of these variables on which \( y \) may depend.

### 2.14. Estimating labor costs for preparing different menu items:
A catering company prepares and serves specialized meals for groups who hire their services to celebrate some occasion or other with a meal party. All the meals include a main course (there are four choices for this: \( M_1 \), chicken; \( M_2 \), meat; \( M_3 \), fish; and \( M_4 \), pasta/vegetarian), an appetizer (there are three choices for this: \( A_1 \), Fruit cocktail; \( A_2 \), Nut and date cocktail; and \( A_3 \), Melon and Prosciutto), a dessert (there are three
choices under this: $D_1$, A piece of pie; $D_2$, A cake or pastry; and $D_3$, Strawberries with ice cream), an optional salad, and optional liqueurs and other after dinner drinks. Of course, all meals include serving of wine or champagne during the main course, and coffee/tea after the main course as standard features.

<table>
<thead>
<tr>
<th>Contract number</th>
<th>Labor cost/ person in $</th>
<th>Choices in agreed menu</th>
<th>Main course</th>
<th>Appetizer</th>
<th>Dessert</th>
<th>Salad</th>
<th>Liqueurs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.65</td>
<td>$M_1$</td>
<td>$A_1$</td>
<td>$D_2$</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>7.25</td>
<td>$M_1$</td>
<td>$A_1$</td>
<td>$D_1$</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6.15</td>
<td>$M_4$</td>
<td>$A_1$</td>
<td>$D_2$</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>6.25</td>
<td>$M_4$</td>
<td>$A_1$</td>
<td>$D_2$</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6.95</td>
<td>$M_1$</td>
<td>$A_1$</td>
<td>$D_1$</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5.75</td>
<td>$M_4$</td>
<td>$A_1$</td>
<td>$D_2$</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7.50</td>
<td>$M_3$</td>
<td>$A_1$</td>
<td>$D_3$</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>7.60</td>
<td>$M_3$</td>
<td>$A_1$</td>
<td>$D_3$</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>7.75</td>
<td>$M_3$</td>
<td>$A_1$</td>
<td>$D_3$</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>5.70</td>
<td>$M_4$</td>
<td>$A_1$</td>
<td>$D_2$</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>6.95</td>
<td>$M_1$</td>
<td>$A_1$</td>
<td>$D_1$</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>6.85</td>
<td>$M_4$</td>
<td>$A_1$</td>
<td>$D_2$</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>7.20</td>
<td>$M_1$</td>
<td>$A_1$</td>
<td>$D_3$</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>7.35</td>
<td>$M_2$</td>
<td>$A_1$</td>
<td>$D_3$</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>7.80</td>
<td>$M_3$</td>
<td>$A_2$</td>
<td>$D_3$</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>8.15</td>
<td>$M_2$</td>
<td>$A_2$</td>
<td>$D_1$</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>9.15</td>
<td>$M_3$</td>
<td>$A_2$</td>
<td>$D_3$</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>7.50</td>
<td>$M_2$</td>
<td>$A_2$</td>
<td>$D_1$</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>6.50</td>
<td>$M_4$</td>
<td>$A_2$</td>
<td>$D_2$</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>8.50</td>
<td>$M_3$</td>
<td>$A_2$</td>
<td>$D_3$</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>7.50</td>
<td>$M_1$</td>
<td>$A_2$</td>
<td>$D_2$</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>8.70</td>
<td>$M_2$</td>
<td>$A_2$</td>
<td>$D_3$</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>9.00</td>
<td>$M_3$</td>
<td>$A_2$</td>
<td>$D_3$</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>7.80</td>
<td>$M_3$</td>
<td>$A_2$</td>
<td>$D_1$</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>7.50</td>
<td>$M_1$</td>
<td>$A_3$</td>
<td>$D_1$</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>6.00</td>
<td>$M_4$</td>
<td>$A_3$</td>
<td>$D_2$</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>8.50</td>
<td>$M_3$</td>
<td>$A_3$</td>
<td>$D_1$</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>7.80</td>
<td>$M_2$</td>
<td>$A_3$</td>
<td>$D_1$</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>7.25</td>
<td>$M_1$</td>
<td>$A_3$</td>
<td>$D_2$</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>6.20</td>
<td>$M_1$</td>
<td>$A_3$</td>
<td>$D_2$</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>5.75</td>
<td>$M_4$</td>
<td>$A_3$</td>
<td>$D_2$</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>7.10</td>
<td>$M_1$</td>
<td>$A_3$</td>
<td>$D_1$</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>7.00</td>
<td>$M_2$</td>
<td>$A_3$</td>
<td>$D_2$</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>7.50</td>
<td>$M_3$</td>
<td>$A_3$</td>
<td>$D_1$</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>6.15</td>
<td>$M_4$</td>
<td>$A_3$</td>
<td>$D_2$</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

1 = yes, 0 = No
The catering company provides from its own store the necessary food and other ingredients for preparing the meals; but hires workers from a sister company for preparing, assembling, and serving the meals. The sister company charges our catering company for the man hours of labor provided to prepare the meals according to the agreed upon menu, serving them, and handling the event. Data on these labor costs from 35 different meal contracts (with 100–150 guests each) is in the table given above.

Develop an LP model to determine the labor costs associated with each choice available under the main course, appetizer, dessert, salad, and liqueurs.

2.15. Estimating the sales potential of a new store at a proposed location: A department store chain wants to develop an estimate of sales for a store as a function of the demographic and other relevant information of the site where it is located.

<table>
<thead>
<tr>
<th>No.</th>
<th>SV</th>
<th>SA</th>
<th>Data related to trading zone of store</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>P</td>
</tr>
<tr>
<td>1</td>
<td>192</td>
<td>122.5</td>
<td>643</td>
</tr>
<tr>
<td>2</td>
<td>236</td>
<td>111.7</td>
<td>743</td>
</tr>
<tr>
<td>3</td>
<td>156</td>
<td>123.4</td>
<td>872</td>
</tr>
<tr>
<td>4</td>
<td>133</td>
<td>106.0</td>
<td>720</td>
</tr>
<tr>
<td>5</td>
<td>102</td>
<td>116.9</td>
<td>491</td>
</tr>
<tr>
<td>6</td>
<td>86</td>
<td>85.0</td>
<td>286</td>
</tr>
<tr>
<td>7</td>
<td>108</td>
<td>90.3</td>
<td>159</td>
</tr>
<tr>
<td>8</td>
<td>74</td>
<td>74.9</td>
<td>190</td>
</tr>
<tr>
<td>9</td>
<td>149</td>
<td>122.8</td>
<td>530</td>
</tr>
<tr>
<td>10</td>
<td>242</td>
<td>120.9</td>
<td>1,113</td>
</tr>
<tr>
<td>11</td>
<td>148</td>
<td>120.3</td>
<td>133</td>
</tr>
<tr>
<td>12</td>
<td>148</td>
<td>64.4</td>
<td>337</td>
</tr>
<tr>
<td>13</td>
<td>117</td>
<td>112.6</td>
<td>309</td>
</tr>
<tr>
<td>14</td>
<td>112</td>
<td>93.8</td>
<td>230</td>
</tr>
<tr>
<td>15</td>
<td>47</td>
<td>49.1</td>
<td>244</td>
</tr>
<tr>
<td>16</td>
<td>110</td>
<td>118.1</td>
<td>280</td>
</tr>
<tr>
<td>17</td>
<td>97</td>
<td>104.5</td>
<td>217</td>
</tr>
<tr>
<td>18</td>
<td>80</td>
<td>110.1</td>
<td>518</td>
</tr>
<tr>
<td>19</td>
<td>90</td>
<td>105.0</td>
<td>202</td>
</tr>
<tr>
<td>20</td>
<td>73</td>
<td>77.2</td>
<td>180</td>
</tr>
<tr>
<td>21</td>
<td>88</td>
<td>85.7</td>
<td>166</td>
</tr>
<tr>
<td>22</td>
<td>132</td>
<td>100.0</td>
<td>647</td>
</tr>
<tr>
<td>23</td>
<td>84</td>
<td>116.1</td>
<td>210</td>
</tr>
<tr>
<td>24</td>
<td>81</td>
<td>74.1</td>
<td>422</td>
</tr>
<tr>
<td>25</td>
<td>134</td>
<td>113.2</td>
<td>607</td>
</tr>
</tbody>
</table>
Each department store has its trading zone, that is, the zone from which it draws its traffic. For a given size of store, all else being equal, the more people in the trading zone, the higher the sales volume at the store. For a typical suburban store, 75% of all customers come from within a 5–10 mile radius; that is why this area is defined as the store’s trading zone. The characteristics of the trading zone like its population \((P\text{ in units of 1,000 people})\), the average family size \((FS)\), the median family income \((MFI \text{ in units of } \$100)\), the selling area in the store \((SA\text{ in } 1,000 \text{ ft}^2)\), the % of hard goods in the store \((HG)\), the percent of home owners \((HO)\), the median home value \((HV \text{ in units of } \$100)\), the median rent paid/month by those living in rented quarters \((RP \text{ in } \$)\) the percent of blacks \((B)\), and hispanics \((H)\) in the population in the trading zone are all expected to influence the sales volume in the store. It is possible to get some of this information using computerized census tract data, but the rest has to be collected.

The chain already has 25 stores in the NY area. Data on the annual sales volume \((SV \text{ in units of } \$100,000)\) at these stores and on the values of these explanatory factors for the sites where they are located is given below (all these data is 1970s data, so the $ refers to 1970s US $). The first column gives the store number.

Develop an LP model to express the expected annual sales at a store in terms of the explanatory variables on which data is provided, that best fits the data.

Solve the model using an available LP software package and find the best fit functional form for the expected annual sales.

There are two potential new sites where the chain has the opportunity to open new stores. These sites are in densely populated areas with good access and little competition, and as there are not enough stores near them, it is believed that residents of those areas are spending significant amounts outside of their zone. Here is the data on the characteristics of the two sites:

**Site 1:** \((P, FS, B, H, MFI, HO, HV, RP) = (955, 3.7, 40.0, 10.8, 84, 10, 23, 80).\)
From the land area available, the selling area in the store set up can be 125,000 \text{ ft}^2, which can be increased to 146,000 \text{ ft}^2 if necessary.

**Site 2:** \((P, FS, B, H, MFI, HO, HV, RP) = (431, 3.5, 13.8, 6.6, 94, 11, 18, 83).\)
and the store can have a selling area of 120,000 \text{ ft}^2.

Use the functional form found above to predict the expected sales volume in a store is set up at sites 1, 2. Which among these sites should get priority for building the next store?

**2.16. (a) Developing a formula for new car sales volume at US franchised dealers by quarter:** The following table gives data on the total volume of sales of new cars at US franchised dealers all over US by quarter from 1958 to 1974.
There is a long-term growth trend in sales due to growth in the population and increases in standard of living. It is believed that the trend factor is of the form $(1069)a^t$, where $a$ is a parameter to be estimated from the data and $t$ is the number of the quarter, counting with 1958, 1st quarter as the one corresponding to $t = 0$.

New car sales are also affected by the season of the year. Let $s_1$ to $s_4$ denote the seasonal factors associated with the seasons corresponding to the first to the fourth quarters of the year.

Another important thing affecting new car sales is the fluctuation in general business activity caused by the so-called odd year–even year business cycle. Let these cyclical factors corresponding to the odd, even years be denoted by $c_1$, $c_2$ respectively. So, our model is

Total new car sales at all franchised dealers in US for a quarter $D$ (the trend factor for the quarter)(seasonal factor for the quarter)(cyclical factor for the year).

Give an LP formulation for the problem of determining the values of various parameters in this model, which gives the best fit for the data.

(b) Developing a formula for monthly champagne sales in France: Historical data on the sales of champagne in France (in millions of bottles) by month for the last six years is given in the following table. French champagne sales have a strong seasonal pattern, but also exhibit a steady growth trend with time.

Using plots of data as necessary develop an appropriate model for (1) French champagne sales in month $j$ of year $t$, in terms of $t$, $j$, and (2) the annual sales in year $t$, in terms of $t$. Formulate the problem of determining the values for the unknown parameters in the model to give the best fit to the data given, as an LP.
### 2.17. Goal programming for project selection:

**IML (Indian Mines Ltd.)** is the largest coal-producing company in India, producing various grades of coal and contributing more than 90% of the country’s production. Being a public sector company, several of the mines operated by it are losing money. The company has decided to face the challenges of (a) increasing total annual production, (b) reducing the average cost of production/ton, and (c) making their mines profitable, or at least reducing their losses from present levels to some reasonable levels.

In this pursuit, the company has decided to invest in two types of mine projects: (1) reconstruction mine projects to enhance production capacity of some existing, operating mines, and (2) opening up new mines. They have made a list of eight different projects of which three are reconstruction projects, and five are new mine projects to invest in. Following table provides data on these projects.

<table>
<thead>
<tr>
<th>Project Number</th>
<th>Total Investment (IN - million rupees)</th>
<th>Present Annual Production (PP - million tons)</th>
<th>Future Estimated Annual Production (FP - million tons)</th>
<th>Present Total Annual Operating Cost (PC - million rupees)</th>
<th>Future Estimated Total Annual Operating Cost (FC - million rupees)</th>
<th>Annual Profit/Loss (PPR - million rupees)</th>
<th>Future Estimated Annual Profit/Loss (FPR - million rupees)</th>
<th>Present Manpower (PM - persons)</th>
<th>Future Estimated Manpower (FM - persons)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.54</td>
<td>3.26</td>
<td>(5, 1)</td>
<td>3.63</td>
<td>7</td>
<td>3.52</td>
<td>7</td>
<td>3.52</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>2.48</td>
<td>1.58</td>
<td>2</td>
<td>4.29</td>
<td>8</td>
<td>1.82</td>
<td>8</td>
<td>1.82</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>3.03</td>
<td>3.53</td>
<td>3</td>
<td>4.15</td>
<td>9</td>
<td>5.22</td>
<td>9</td>
<td>5.22</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>3.27</td>
<td>5.21</td>
<td>4</td>
<td>4.12</td>
<td>10</td>
<td>6.87</td>
<td>10</td>
<td>6.87</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>3.78</td>
<td>7.62</td>
<td>5</td>
<td>4.65</td>
<td>11</td>
<td>10.80</td>
<td>11</td>
<td>10.80</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>3.23</td>
<td>9.25</td>
<td>6</td>
<td>4.75</td>
<td>12</td>
<td>13.92</td>
<td>12</td>
<td>13.92</td>
<td>12</td>
</tr>
</tbody>
</table>
For simplicity assume that on each project \( j \), the company may decide to invest a fraction \( x_j \) of the total investment needed for it, where \( 0 \leq x_j \leq 1 \) for all \( j \) and \( x = (x_j) \). Make the simplifying linearity assumption that the effect of \( x \) for an existing mine \( j \) will be: values after the decision is implemented, of: annual production, annual operating cost, annual profit/loss, manpower, will be: (its present value)(1 \( - x_j \)) + \( x_j \) (future estimated value mentioned in the tableau); and for a new mine \( j \) these values will be: \( x_j \) (future estimated value mentioned in the tableau).

A team of managers at the company has selected the following as important targets or goals for the investment decision: (1) if possible, keep total investment in all projects to within 1,100 million rupees, (2) try to achieve an average production cost of 350 rupees/ton of coal produced at these mines, (3) try to keep the total loss at these project sites to within 60 million rupees as far as possible, and (4) keep the total manpower deployment at these project sites ≤ the desired upper limit of 9,000.

Also, after consulting several executives and experts, it has been determined that the weights measuring the achievements of various goals should be 0.298 for the capital investment goal, 0.252 for the production cost goal, 0.241 for the profit/loss goal, and 0.209 for the manpower goal.

The company would like to impose a constraint that the total estimated annual production at these sites should reach or exceed 2 million tons.

(1) Formulate the problem as a linear goal programming problem.

(2) Discuss what changes should be made in the model if it is required that each project should either be not taken up at all; or if it is taken up, it should be completed (i.e., partial funding of projects is not allowed) (Mukherjee and Bera 1995).

2.18. A bicriterion assignment problem: Let \( K \) denote the set of feasible solutions of the usual \( n \times n \) assignment problem represented by the system of constraints: \( \{ x = (x_{ij}) : \sum_{i=1}^{n} x_{ij} = 1 \) for all \( j = 1 \) to \( n \); \( \sum_{j=1}^{n} x_{ij} = 1 \) for all \( i = 1 \) to \( n \}. \)

In the usual cost-minimizing linear assignment problem (LAP), the decision variable \( x_{ij} \) takes the value 1 [0] if job \( i \) is assigned [not assigned] to machine \( j \) for
execution; \(c_{ij}\) represents the job completion cost of carrying out job \(i\) on machine \(j\), and the objective function to minimize is \(z_1(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij}\).

All extreme point solutions of \(K\) (called assignments) make all variables \(x_{ij}\) take only values 0 or 1. That is why all algorithms for LAP try to obtain extreme point optimum solutions, thereby meeting the 0-1 requirements on all the variables without imposing these requirements explicitly.

Let \(t_{ij}\) denote the time (in hours, say) needed to complete job \(i\) on machine \(j\). Then the time taken to complete all jobs if assignment \(x^* = (\tilde{x}_{ij})\) is implemented is \(\tilde{t} = \max\{t_{ij} : (i,j)\) such that \(\tilde{x}_{ij} = 1\}\), assuming that all the jobs are started simultaneously.

Normally when work is going on, the company needs to have some supervisors to supervise the work. Suppose the cost of providing the supervisory force is \(F/s.h\). Both the job completion cost and the supervisory cost need to be minimized, and so this leads to a bicriterion assignment problem: minimize \(z_1(x)\) and \(z_2(x) = F \max\{t_{ij} : x_{ij} = 1\}\) subject to \(x \in K\) and \(x_{ij} \in \{0,1\}\) for all \(i,j\).

(1) Prove that this bicriterion assignment problem has at most \(n^2\) efficient solutions.
(2) Prove that an optimum solution for the single objective problem of minimizing \(z_1(x) + z_2(x)\) over \(x \in K\) is attained at an efficient solution of the bicriterion assignment problem.

Clearly, \(z_2(x)\) is not a linear function of \(x\). Consider the following algorithm for the problem of minimizing \(Q(x) = z_1(x) + z_2(x)\) (not a linear function) over \(x \in K\). Let \(\ell\) be a lower bound for \(z_2(x)\) over \(x \in K\). For example, you can arrange \(t_{ij}\)s in increasing order in a list, and take \(\ell\) to be the \(n\)th element from the bottom in this ordered list. Let \(c^0 = (c^0_{ij} = c_{ij})\).

Find an optimum assignment, \(x^0\), for the LAP of minimizing \(\sum_{i=1}^{n} \sum_{j=1}^{n} c^0_{ij} x_{ij}\). Let \(z^0_1 = z_1(x^0), z^0_2 = z_2(x^0), Q^0 = z^0_1 + z^0_2\). Define \(x^* = x^0, Q^* = Q^0\). Set \(k = 1\) and go to Step 1.

**General Step \(k\):** Consists of two substeps.

**Substep 1:** Define for all \(i,j\), \(c^k_{ij} = c_{ij}\) if \(t_{ij} < z^{k-1}_2\), or \(= \infty\) if \(t_{ij} \geq z^{k-1}_2\).

**Substep 2:** Find an optimum assignment \(x^k\) minimizing \(\sum_{i=1}^{n} \sum_{j=1}^{n} c^k_{ij} x_{ij}\). Let \(z^k_1 = z_1(x^k), z^k_2 = z_2(x^k), Q^k = z^k_1 + z^k_2\).

If \(z^k_1 \geq Q^* - \ell\), then \(x^*\) is an optimum solution of the problem and \(Q^*\) is the optimum objective value, terminate.

If \(Q^k < Q^*\), update \(x^*\) to \(x^k\), \(Q^*\) to \(Q^k\), and go to the next step by adding 1 to \(k\).

In this algorithm show that \(z^k_1\) is nondecreasing, while \(z^k_2\) is strictly decreasing. Using this show that the algorithm obtains an optimum solution minimizing \(Q(x)\) over \(K\). When

\[
\begin{array}{cccccc}
6 & 3 & 5 & 8 & 10 & 6 \\
6 & 4 & 6 & 5 & 9 & 8 \\
11 & 7 & 4 & 8 & 3 & 2 \\
9 & 10 & 8 & 6 & 10 & 4 \\
4 & 6 & 7 & 9 & 8 & 7 \\
3 & 5 & 11 & 10 & 12 & 8 \\
\end{array}
\]

and

\[
\begin{array}{cccccc}
4 & 20 & 9 & 3 & 8 & 9 \\
6 & 18 & 8 & 7 & 17 & 8 \\
2 & 8 & 20 & 7 & 15 & 7 \\
12 & 13 & 14 & 6 & 9 & 10 \\
9 & 8 & 7 & 14 & 5 & 9 \\
17 & 13 & 3 & 4 & 13 & 7 \\
\end{array}
\]


find an optimum assignment minimizing $Q(x)$ \cite{Geetha93, Bakshi79, Berman90}.

### 2.19. Planning diets for diabetic patients:

An important part of managing diabetes is very careful diet planning and sticking to it conscientiously. Diabetic patients have to observe recommended dietary allowances on carbohydrates, fats, and calories; and also make sure that their daily diet consists of other nutrients such as vitamin B$_6$, iron, etc., within prescribed limits.

<table>
<thead>
<tr>
<th>Nutrient</th>
<th>Average units/serving of food group</th>
<th>RDA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MI</td>
<td>VEG</td>
</tr>
<tr>
<td>CL</td>
<td>118</td>
<td>20</td>
</tr>
<tr>
<td>PR</td>
<td>9.3</td>
<td>1.4</td>
</tr>
<tr>
<td>CH</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>F</td>
<td>3.1</td>
<td>0.2</td>
</tr>
<tr>
<td>CA</td>
<td>334</td>
<td>44</td>
</tr>
<tr>
<td>FE</td>
<td>0.15</td>
<td>0.75</td>
</tr>
<tr>
<td>VC</td>
<td>2.3</td>
<td>22.84</td>
</tr>
<tr>
<td>VB6</td>
<td>90.2</td>
<td>91.09</td>
</tr>
<tr>
<td>VB12</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>VE</td>
<td>0.05</td>
<td>1.51</td>
</tr>
<tr>
<td>MG</td>
<td>28.8</td>
<td>21.6</td>
</tr>
<tr>
<td>ZN</td>
<td>1.2</td>
<td>0.4</td>
</tr>
</tbody>
</table>

The usual approach to menu planning is a manual one based on exchange lists of food groups, but most patients find this tedious and difficult to use. This exercise deals with menu planning considering both client food preferences and recommended dietary allowances, using an interactive goal programming approach.

In their exchange lists for meal planning, the ADA (American Diabetes Association) forms various foods into six food groups: MI (milk), VEG (vegetables), FT (fruits), BR (breads, this group also includes crackers, cereal, starchy vegetables), ME (meats), OI (oils). For each food group, 1 serving is defined to be the following quantity: MI (1 cup or 245 g), VEG (1 cup cooked vegetables or 80 g), FT (averaged value of 100 g), BR (1 slice or 25 g for bread and/or crackers, 56 g for cereal, or 118 g for starchy vegetables), ME (1 oz or 30 g of lean meat), OI (1 teaspoon or 14 g of oil, butter, or margarine).

For each sex and age group of the patient, ADA has provided RDA (recommended dietary allowance) for various nutrients. Among all the nutrients, experimentation has indicated that meeting the RDAs for a selected list of 12 generally produced a diet sufficient also in the others. These 12 nutrients are CL (calories), PR (proteins, measured in grams), CH (carbohydrates, measured in grams), F (fat, measured in grams), CA (calcium, measured in milligrams), FE (iron, measured in micrograms), VC (vitamin C, measured in milligrams), VB6 (vitamin B$_6$, measured
in milligrams), VB12 (vitamin $\text{B}_{12}$, measured in micrograms), VE (vitamin E, measured in IU), MG (magnesium, measured in milligrams), and ZN (zinc, measured in milligrams).

The patient is a young adult female. For her sex and age group, the RDA for various nutrients are given in the last column of the table given below. The table also gives the average value of the nutrient content/serving of each food group, obtained using the Ohio State University. Nutrient Data Bank. She has specified the minimum and maximum number of servings of each food group that she would like to include in her daily diet. These bounds are MI (2–5 servings), VEG (1–5 servings), FT (1–5 servings), BR (3–6 servings), and OI (2–5 servings). Also, the ADA has recommended that no more than 35% of calories in the diet should be derived from the MI and OI groups. All these define the constraints on the decision variables in the model for the problem.

The RDA number for each of the 12 nutrients is the goal for the content of that nutrient in the daily diet. The aim is to find a diet in which the contents of the nutrients CL, CH, and F are less than the corresponding goals as far as possible, while the contents of the other nutrients are as close to the goal as possible. Take the weight (a measure of the importance attached to that goal achievement) for goals on CL, CH, and F as 5 each; the weight for each of the vitamin goals as 4; and the weight for all the other goals as 3. Formulate the problem of determining the number of servings of each of the food groups to include in the daily diet, to meet the goals subject to the constraints mentioned above (Rugg et al. 1983).

2.20. Job allocation to operators at a book typesetting company: An agency typesets books for publishers. At present they are starting work on $m$ different books, and $n$ operators are available to work on typesetting them. The operators work independently, each at his own speed, but all of them begin work on the portions assigned to them at time point 0 and continue until it is finished. The agency can split each book into any number of fractions, so that different fractions can be performed by different operators.

If $x_{ij}$ denotes the fraction of the $i$th book allotted to the $j$th operator, from past experience they can estimate that it will take that operator $\gamma_{ij} x_{ij}$ units of time, where the $m \times n$ matrix $\Gamma = (\gamma_{ij})$ is given. The total working time of an operator will be the sum of the times needed to perform all activity fractions assigned to him/her.

Formulate the problem of determining the fractions of each job to be allocated to each operator, so as to minimize the clock time by which all the jobs are completed, as an LP. Give this formulation for the numerical example in which $m = 5$, $n = 3$, and

$$
\Gamma = \begin{pmatrix}
13 & 17 & 7 \\
37 & 23 & 11 \\
19 & 2 & 31 \\
5 & 3 & 29 \\
1 & 1 & 1
\end{pmatrix}.
$$

(Andreatta et al. 1993).
2.21. Tele-marketing: A market research center (MRC) needs to collect data on three types of products. They hire hourly employees to call potential customers and gather this information. During a certain hourly period, the number of calls to be made regarding the three types of products is given in the following table. Here expected call duration is in minutes (min).

<table>
<thead>
<tr>
<th>Product type</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. calls to make</td>
<td>90</td>
<td>150</td>
<td>270</td>
</tr>
<tr>
<td>Expected call duration</td>
<td>2.3</td>
<td>3.4</td>
<td>1.9</td>
</tr>
</tbody>
</table>

(1) There are six grades of employees with different salaries (wages/hour given in dollars) for making these calls. Each employee can handle calls on a subset of products depending on their expertise and past experience. Here is the relevant data.

<table>
<thead>
<tr>
<th>Call types they handle</th>
<th>Grade $\rightarrow$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1</td>
<td></td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Type 2</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Type 3</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Wage/hour</td>
<td></td>
<td>10</td>
<td>8</td>
<td>25</td>
<td>20</td>
<td>22</td>
<td>30</td>
</tr>
</tbody>
</table>

"*" indicates grade can handle this type

Let $x_j$ = number of employees of grade $j$ employed during the hour for making these calls. Taking $x_j$ as a continuous variable, model the problem of getting the job done at the smallest cost.

(2) Now assume that they have a 3 h period to complete the above job, but a limit of only 14 employees who can work in the office at any point of time. They have the opportunity to hire the employees for a period of 1 h, or 2 h, or 3 h for this work. Those hired for 1 h are to be paid at the hourly rate given above. Those hired for a 2 h period are paid at the above rate for the first hour, and 90% of the above rate for the second hour. Those hired for the 3 h period are to be paid for the first two hours as mentioned in previous sentence, but only 80% of the hourly rate given in the above table for the third hour.

Discuss the model to find the best plan to hire people to complete the job at minimal cost in this situation. (Problem given by Vincent F. Yu.)

2.22. Minimizing losses from attacks on bank armored vehicles: In large cities, each bank usually has several branches scattered in different areas of the city. On each working day, the need arises for the bank to transfer different quantities of money between its branches in the city. Banks use armored vehicles for these money transfers. While these armored vehicles are traveling on the roads, they are sometimes attacked by robbers to loot the money inside them.

Designing safe routes for their armored vehicles is a serious problem that banks face. The chance of an attack on these bank vehicles is usually very low, but it
increases in hard economic times as unemployment in the city goes up. When an
attack occurs, if the attackers are successful, the bank stands to lose all the money
inside the vehicle at that time. One strategy that banks can use to minimize the
expected loss from these attacks is to put an upper bound, \( u \), on the amount of
money that can be carried inside an armored vehicle at any time.

As the upper bound \( u \) decreases, the expected total daily mileage, \( m \), of all the
armored vehicles of the bank in the city increases as these vehicles may have to
make more trips to carry out the needed money transfers. So, let \( m(u) \) denote this
expected total daily mileage of the banks armored vehicles as a function of \( u \).

Using simulation, the functional form of \( m(u) \) has been determined to be the
following PL function. Here, \( u, m(u) \) are given in coded money, distance units.

\[
\begin{array}{|c|c|}
\hline
\text{Range of } u & m(u) \\
\hline
40–50 & 110 - 3(u - 50) \\
35–40 & 140 - 5(u - 40) \\
30–35 & 165 - 10(u - 35) \\
\hline
\end{array}
\]

As the expected daily mileage increases, their public exposure increases, and
this is expected to increase the chance of an attack on them. Let \( p(m) \) denote the
probability of an attack on one of these armored vehicles while on the road in a
day, as a function of \( m \). Using data on the past records of these occurrences and
other information, the Statistics Division of the bank came up with the following
estimates of \( p(m) \).

\[
\begin{array}{|c|c|}
\hline
\text{Range of } m & \text{Estimate of } p(m) \\
\hline
110 – 155 & 0.0018 + 0.0005((m - 110)/45) \\
155 – 215 & 0.0023 + 0.0007((m - 155)/60) \\
\hline
\end{array}
\]

From past data, we know that once an attack occurs on one of their armored
vehicles on the road, the probability that the attack will be repelled with no loss of
money inside is 0.25 and the probability that all the money inside will be lost is
0.75.

Using this information determine the optimum value of \( u \), explaining clearly how
it is obtained.

(1) With fuel prices going up, extra mileage increases the cost of armored vehicle
operations too. Discuss how the model will change if it is required to minimize
the sum of expected losses from attacks on armored vehicles plus the expected
cost of armored vehicle operations to carry out the needed money transfers.

2.23. Referring to Sect. 2.5, prove that if \( \tilde{x} \) is an optimum solution of (2.16)
minimizing a positive weighted combination of all the objective functions in a multi-
objective minimization problem (2.15), then \( \tilde{x} \) is a nondominated solution for (2.15).

2.24. Concave function minimization: Consider the problem: Minimize \( \theta(x) \) subject
to \( Ax = b, \ x \geq 0 \), where \( A \) is a matrix of order \( m \times n \) and \( \theta(x) \) is a concave
function. If this problem has an optimum solution, show that it must have an optimum solution which is an extreme point of the set of all feasible solutions.

2.25. Minimizing scrap in aluminium foil slitting: At an aluminium foil mill, the input is in the form of coils of aluminium foil in standard widths (SW), which they slit into foil of different widths for a variety of end users. Customers order foil specifying the width, gauge, and surface finish required. The mill selects an appropriate coil from stock and slits it into the desired width. On a particular day they have standard widths A to G in stock, and need to fill customer widths (CW) I to V. Following table gives the weight of scrap in pounds, which will result from slitting the customer orders of each width on the leftmost column of the table, from the SW on the top line (an entry of “.” indicates that the SW is not suitable for that CW). We need to determine which SW to use in slitting each CW to minimize total scrap generated for filling all customer orders. Can this problem be modeled as an LP? Formulate this problem. (Lanzenauer 1975).

<table>
<thead>
<tr>
<th>CW</th>
<th>SW</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>61</td>
<td>19</td>
<td>69</td>
<td>4</td>
<td>46</td>
<td>26</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>15</td>
<td>.</td>
<td>44</td>
<td>52</td>
<td>66</td>
<td>95</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>94</td>
<td>55</td>
<td>.</td>
<td>85</td>
<td>65</td>
<td>67</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>42</td>
<td>48</td>
<td>11</td>
<td>62</td>
<td>13</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>23</td>
<td>.</td>
<td>.</td>
<td>58</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
</tbody>
</table>

2.26. We need to find a vector \( x = (x_1, x_2)^T \) satisfying each of the following equations as closely as possible: \( (2x_1 + 4x_2, x_1 + 2x_2, 4x_1 + 5x_2, 4x_1 + 3x_2, 3x_1 + 4x_2, 6x_1 + 3x_2)^T = (240, 130, 320, 170, 60, 45)^T \). Formulate this as an LP and solve it.

2.27. We need to find a vector \( x = (x_1, x_2, x_3)^T \geq 0 \) satisfying the following constraints as closely as possible. Formulate this as an LP.

\[
\begin{align*}
x_1 + x_2 + x_3 &= 100, \\
x_1 + x_2 - x_3 &\geq 60, \\
-x_1 + x_2 + x_3 &\geq 70, \\
-x_1 + x_2 - x_3 &\leq 20.
\end{align*}
\]

2.28. Yearly production plan at a brewery: The demand from retailers for the brewed product at a brewery is highly seasonal. For the months of January to December for the planning year, the demand in units (1 unit = 1,000 barrels) is estimated to be \( (5, 4, 5, 7, 10, 12, 14, 13, 11, 8, 10, 7)^T \).

Each month of the planning year, the brewery can operate at four production levels; one shift/day, two shifts/day, or each with overtime. Following table provides the production data for each of these levels. Money is measured in units, which we will abbreviate as \( mu \).
2 Formulation Techniques Involving Transformations of Variables

<table>
<thead>
<tr>
<th>Prod. level</th>
<th>Prod. capacity (units/month)</th>
<th>Labor cost mu/month</th>
</tr>
</thead>
<tbody>
<tr>
<td>One shift</td>
<td>5</td>
<td>3000</td>
</tr>
<tr>
<td>One shift + overtime</td>
<td>7</td>
<td>4000</td>
</tr>
<tr>
<td>Two shifts</td>
<td>9</td>
<td>6000</td>
</tr>
<tr>
<td>Two shifts + overtime</td>
<td>12</td>
<td>8000</td>
</tr>
</tbody>
</table>

Increasing from one shift level in a month to two shift level the next month incurs a cost of 1,500 mu; while decreasing from two shift level to one shift level incurs a cost of 1,000 mu.

They can hold brewed product between the brewhouse and the bottleshop in refrigerated tanks. Every unit inventory held in these tanks at the end of a production month to the next month costs 300 mu.

Shortage in a month incurs a shortage cost of 500 mu/unit short; also this shortage has to be made up in succeeding months in order to meet total demand. January begins with two units of brewed product in inventory, and would be operating at the one shift production level. Formulate the problem of developing a minimum cost production plan for the planning year that would result in a closing inventory at the end of the year of two units. (From Lanzenauer (1975)).

2.29. Toy-store problem: A toy-store chain has several stores in the midwest. For the coming X-mas season, they need to place orders with their overseas suppliers before the end of May for delivery in time for the X-mas sales period.

As unsold toys at the end of the X-mas season do not contribute much to the profit of the company, they base their order quantities quite close to the expected sales volume. From experience over the years they observed that the X-mas sales volume has a high positive correlation with the DJA = Dow Jones average (a measure of the economic status of the region prior to the sales period), and a high negative correlation with the percent unemployment rate in the region. Following table gives data on the DJA during the months of February, March, April (these are independent variables $x_1$, $x_2$, $x_3$), the percent unemployment in the region during this period (independent variable $x_4$), and the toy sales volume in the region in millions of dollars during the X-mas sales season (dependent variable $y$) between 1990 and 2001.

From the above discussion it is reasonable to assume that the expected value of $y$ can be approximated by a function $a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4$, where the parameters satisfy $0 \leq a_1 \leq a_2 \leq a_3$ and $a_4 \leq 0$. Write the LP formulation of the problem of finding parameter values that give the closest fit to data by the $L_1$ measure of deviation.
## 2.30. Oil Refinery Optimization

A refinery has a distillation capacity of 100,000 barrels of crude/day in its fractionator. Here crude oil is basically heated, and as the temperature increases, different products called DN (distillation naphtha), DHO (distillation heating oil), DGO (distillation gas oil), and P (pitch) are given off in vapor form and are collected at various levels. The refinery gets crude oil from three different countries, these are called crudes 1, 2, 3. All the crudes and the various products are measured by volume in barrels. The output statistics from the distillation of each of the available crudes are tabulated below.

<table>
<thead>
<tr>
<th>Year</th>
<th>Crude 1</th>
<th>Crude 2</th>
<th>Crude 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>10,690</td>
<td>10,185</td>
<td>10,306</td>
</tr>
<tr>
<td>2000</td>
<td>10,533</td>
<td>10,525</td>
<td>10,798</td>
</tr>
<tr>
<td>1999</td>
<td>9,356</td>
<td>9,550</td>
<td>10,307</td>
</tr>
<tr>
<td>1998</td>
<td>8,307</td>
<td>8,664</td>
<td>8,940</td>
</tr>
<tr>
<td>1997</td>
<td>6,828</td>
<td>6,727</td>
<td>6,800</td>
</tr>
<tr>
<td>1996</td>
<td>5,435</td>
<td>5,527</td>
<td>5,579</td>
</tr>
<tr>
<td>1995</td>
<td>3,927</td>
<td>4,080</td>
<td>4,239</td>
</tr>
<tr>
<td>1994</td>
<td>3,898</td>
<td>3,723</td>
<td>3,634</td>
</tr>
<tr>
<td>1993</td>
<td>3,344</td>
<td>3,405</td>
<td>3,434</td>
</tr>
<tr>
<td>1992</td>
<td>3,247</td>
<td>3,253</td>
<td>3,294</td>
</tr>
<tr>
<td>1991</td>
<td>2,798</td>
<td>2,903</td>
<td>2,895</td>
</tr>
<tr>
<td>1990</td>
<td>2,607</td>
<td>2,665</td>
<td>2,673</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>x₁</th>
<th>x₂</th>
<th>x₃</th>
<th>x₄</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>10,690</td>
<td>10,185</td>
<td>10,306</td>
<td>4.3</td>
<td>59</td>
</tr>
<tr>
<td>2000</td>
<td>10,533</td>
<td>10,525</td>
<td>10,798</td>
<td>4.0</td>
<td>60</td>
</tr>
<tr>
<td>1999</td>
<td>9,356</td>
<td>9,550</td>
<td>10,307</td>
<td>4.3</td>
<td>54</td>
</tr>
<tr>
<td>1998</td>
<td>8,307</td>
<td>8,664</td>
<td>8,940</td>
<td>4.6</td>
<td>47</td>
</tr>
<tr>
<td>1997</td>
<td>6,828</td>
<td>6,727</td>
<td>6,800</td>
<td>5.2</td>
<td>36</td>
</tr>
<tr>
<td>1996</td>
<td>5,435</td>
<td>5,527</td>
<td>5,579</td>
<td>5.5</td>
<td>28</td>
</tr>
<tr>
<td>1995</td>
<td>3,927</td>
<td>4,080</td>
<td>4,239</td>
<td>5.5</td>
<td>20</td>
</tr>
<tr>
<td>1994</td>
<td>3,898</td>
<td>3,723</td>
<td>3,634</td>
<td>6.5</td>
<td>17</td>
</tr>
<tr>
<td>1993</td>
<td>3,344</td>
<td>3,405</td>
<td>3,434</td>
<td>7.1</td>
<td>14</td>
</tr>
<tr>
<td>1992</td>
<td>3,247</td>
<td>3,253</td>
<td>3,294</td>
<td>7.4</td>
<td>13</td>
</tr>
<tr>
<td>1991</td>
<td>2,798</td>
<td>2,903</td>
<td>2,895</td>
<td>6.6</td>
<td>11</td>
</tr>
<tr>
<td>1990</td>
<td>2,607</td>
<td>2,665</td>
<td>2,673</td>
<td>5.2</td>
<td>10</td>
</tr>
</tbody>
</table>

### Price ($/barrel) and Available (barrels/day)

<table>
<thead>
<tr>
<th></th>
<th>DN</th>
<th>DHO</th>
<th>DGO</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crude 1</td>
<td>0.19</td>
<td>0.27</td>
<td>0.38</td>
<td>0.05</td>
</tr>
<tr>
<td>Crude 2</td>
<td>0.16</td>
<td>0.32</td>
<td>0.27</td>
<td>0.13</td>
</tr>
<tr>
<td>Crude 3</td>
<td>0.02</td>
<td>0.24</td>
<td>0.26</td>
<td>0.39</td>
</tr>
</tbody>
</table>

| Price ($)/barrel | 23.25 | 22.00 | 20.50 |
| Available (barrels/day) | 60,000 | 90,000 | 80,000 |

It can be verified that the total volume of outputs from the distillation of one barrel of crude is <1. The loss is due to evaporation and unusable heavy residuals.

DHO can be sold directly as heating oil. DGO can be sold directly as diesel fuel. Sale prices of these products are given below.

DHO and DGO can also be processed further in a catalytic cracker. The catalytic cracker can either process a maximum of 100,000 barrels/day of DHO, or a maximum of 50,000 barrels/day of DGO, or a combination of these in proportion of these levels adding up to 1. Also, when processing DHO, the catalytic cracker can be run
either at a normal level or at a high severity level. The high severity level helps to convert more of the DHO into naptha as seen from the table below. In processing DGO, the catalytic cracker is run at normal level only and never on high severity level.

<table>
<thead>
<tr>
<th>Catalytic cracker outputs</th>
<th>DHO normal level</th>
<th>DHO high severity level</th>
<th>DGO normal level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catalytic naptha (CN)</td>
<td>0.18</td>
<td>0.32</td>
<td>0.48</td>
</tr>
<tr>
<td>Catalytic heating oil (CHO)</td>
<td>0.80</td>
<td>0.69</td>
<td>0.52</td>
</tr>
<tr>
<td>Pitch (P)</td>
<td>0.11</td>
<td>0.10</td>
<td>0.15</td>
</tr>
</tbody>
</table>

The cracking process converts the feed into products whose density is smaller than that of the feed, that is why the volume of outputs from this process is greater than the feed volume.

The pitch (from fractionator and catalytic cracker) can be combined with CHO (two parts of CHO to 17 parts of pitch) and sold as heavy fuel oil. DN and CN can be combined (20 parts of DN with 17 parts or greater of CN) and sold as gasoline. The quality of the gasoline improves with the proportion of CN in this blend. The following table gives the selling prices (all money figures in this exercise are in 1995 US $) and demand for the various final products.

<table>
<thead>
<tr>
<th>Final product</th>
<th>Selling price ($/barrel)</th>
<th>Daily demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gasoline</td>
<td>43.25</td>
<td>Up to 40,000 barrels</td>
</tr>
<tr>
<td>Heating oil</td>
<td>39.75</td>
<td>Up to 40,000 barrels</td>
</tr>
<tr>
<td>Diesel fuel</td>
<td>39.00</td>
<td>Any amount</td>
</tr>
<tr>
<td>Heavy fuel oil</td>
<td>30.00</td>
<td>Any amount</td>
</tr>
</tbody>
</table>

The processing cost on the fractionator is estimated to be $0.60/barrel of crude processed. On the catalytic cracker, the processing costs are $1.50/barrel of fresh feed at the high severity level and $0.95/barrel of fresh feed at the normal level. Formulate the problem of determining how much of each final product to produce daily in order to maximize daily net profit, as an LP.

2.31. Consider the following optimization problem in \( n \) variables \( x_1, \ldots, x_n \). In this problem, \( a, h \) are positive data elements.

Minimize \( z(x) = cx \)

subject to \( Dx \geq d \)

\[
\sum_{j=1}^{n} |x_j - a| \leq h
\]

\( x \geq 0 \).
Notice that there is one constraint here involving absolute value terms. Is it possible to transform this problem into a linear program? Explain clearly. Discuss how one can handle this problem.

2.32. A UM Dental School statistician has conducted research on the condition of the teeth of 8-year-old kids in Michigan and estimated that the average number of decayed teeth in the mouth of an 8-year-old Michigan kid is 1.8.

Given this data, it is required to find the minimum and maximum possible values for the fraction of 8-year-old Michigan kids who have two or more decayed teeth in their mouths. Give LP formulations for the problems of finding this minimum and maximum (two separate problems).

Also comment on why an LP model is appropriate for these problems.

2.33. Optimum assignment of students to schools: A school district consists of \( r \) neighborhoods, \( s \) schools, and in each of these schools \( g \) is the number of grades or levels (like fourth grade, fifth grade, etc.) for students. Here are the data elements: \( k_{ju} \) = capacity for the number of students of grade \( u \) in school \( j \); \( s_{iu} \) = number of students in neighborhood \( i \) studying in grade \( u \) in a particular school year; and \( d_{ij} \) = minutes that a student from neighborhood \( i \) has to spend in school bus daily to get to school \( j \), for \( j = 1 \) to \( s \), \( u = 1 \) to \( g \), and \( i = 1 \) to \( r \).

Treating the number of students as a continuous variable, formulate the problem of assigning students in this school district to schools to minimize the total number of minutes spent in the school bus daily by all the students as a linear program. Also, discuss whether this problem can be formulated as a transportation problem.

2.34. Recruiting workers for a new plant: A paper company is setting up a new plant in a city for which they need to recruit the following numbers of workers: 2,000 nonprofessional workers and 800 professional workers.

There are four categories of workers: men (nonminority), women (nonminority), men (minority), women (minority). The estimated average cost of recruiting per worker in each category is given below:

<table>
<thead>
<tr>
<th>Category</th>
<th>Nonprofessional</th>
<th>Professional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minority (men or women)</td>
<td>$740</td>
<td>$1560</td>
</tr>
<tr>
<td>Women (nonminority)</td>
<td>850</td>
<td>1450</td>
</tr>
<tr>
<td>Men (nonminority)</td>
<td>570</td>
<td>1290</td>
</tr>
</tbody>
</table>

The company has established the following goals in order of priority:

**Goal 1:** The percentage of women (minority or nonminority) among the workers should be at least 30% as far as possible.

**Goal 2:** The percentage of minorities (men or women) among the workers should be at least 20% as far as possible.

**Goal 3:** The recruiting cost should be at most $2 million as far as possible.

It is required to find how many workers to recruit in each category in order to meet the goals as closely as possible. Ignoring the integer nature of the variables, formulate this problem as a linear program using a goal programming approach.

If you need to make any additional assumptions, please state them very clearly.
2.35. **Short-term cash flow management:** A department store has revenues coming from their operations, and bills to pay out, at the end of each month. These quantities for a 6-month period are given below.

<table>
<thead>
<tr>
<th>Month</th>
<th>Revenue from operations</th>
<th>Bills to pay</th>
</tr>
</thead>
<tbody>
<tr>
<td>July</td>
<td>$1,000</td>
<td>5,000</td>
</tr>
<tr>
<td>Aug</td>
<td>2,000</td>
<td>5,000</td>
</tr>
<tr>
<td>Sep</td>
<td>2,000</td>
<td>6,000</td>
</tr>
<tr>
<td>Oct</td>
<td>4,000</td>
<td>2,000</td>
</tr>
<tr>
<td>Nov</td>
<td>7,000</td>
<td>2,000</td>
</tr>
<tr>
<td>Dec</td>
<td>9,000</td>
<td>1,000</td>
</tr>
</tbody>
</table>

Clearly, they have a short-term cash flow problem at the beginning of the period, as they have no other income to pay the bills other than the revenue. They have two available sources to borrow money.

**Bank:** The bank can loan a maximum of $7,000. The loan can be taken on 31 July, and it should be paid back together with 9% interest on 31 Dec (early payback does not reduce the interest amount).

**S&L:** The S&L can loan any amount of money for a period of one month, any number of times. The loan amount can be taken on the last day of any month, and it should be paid back together with 3% interest on the last day of following month.

Formulate the problem of determining how the store can minimize the cost of paying their bills on time each month in the period, as a linear program.

2.36. **Optimizing currency exchange transactions:** A company holds 1.2 billion Japanese yen, 10.5 billion Indonesian rupiahs, and 28 million Malaysian ringgits. Here are the exchange rates and transaction costs. JY = Japanese yen, IR = Indonesian rupiah, MR = Malaysian ringgit, UD = US$, CD = Canadian$, EE = European euro, EP = English pound, MP = Mexican peso.

<table>
<thead>
<tr>
<th>From</th>
<th>JY</th>
<th>IR</th>
<th>MR</th>
<th>UD</th>
<th>CD</th>
<th>EE</th>
<th>EP</th>
<th>MP</th>
</tr>
</thead>
<tbody>
<tr>
<td>JY</td>
<td>1</td>
<td>50</td>
<td>0.04</td>
<td>0.008</td>
<td>0.01</td>
<td>0.0064</td>
<td>0.0048</td>
<td>0.0768</td>
</tr>
<tr>
<td>IR</td>
<td>1</td>
<td>0.0008</td>
<td>0.00016</td>
<td>0.0002</td>
<td>0.000128</td>
<td>0.000096</td>
<td>0.001536</td>
<td></td>
</tr>
<tr>
<td>MR</td>
<td>1</td>
<td>0.2</td>
<td>0.25</td>
<td>0.16</td>
<td>0.6</td>
<td>9.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UD</td>
<td>1</td>
<td>1.25</td>
<td>0.8</td>
<td>0.6</td>
<td>0.12</td>
<td>1.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD</td>
<td>1</td>
<td>0.64</td>
<td>0.48</td>
<td>7.68</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EE</td>
<td>1</td>
<td>0.75</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EP</td>
<td>1</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MP</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.6 Exercises

From Transaction cost %

<table>
<thead>
<tr>
<th></th>
<th>JY</th>
<th>IR</th>
<th>MR</th>
<th>UD</th>
<th>CD</th>
<th>EE</th>
<th>EP</th>
<th>MP</th>
</tr>
</thead>
<tbody>
<tr>
<td>JY</td>
<td>0.5</td>
<td>0.5</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.25</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>IR</td>
<td>0.7</td>
<td>0.5</td>
<td>0.3</td>
<td>0.3</td>
<td>0.75</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MR</td>
<td>0.7</td>
<td>0.7</td>
<td>0.4</td>
<td>0.45</td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UD</td>
<td>0.15</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD</td>
<td></td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EE</td>
<td></td>
<td></td>
<td></td>
<td>0.05</td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>MP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(1) The portion below the main diagonal in the first table is left blank because each entry in it can be obtained from the corresponding entry above the diagonal (every currency considered can be converted into every other currency). The second table is symmetric around the main diagonal, that is why only the portion above the main diagonal is given. The transaction cost for converting one currency, say $c_1$ units of currency 1, into any other currency is $(c_1/100)\times (entry$ $for$ $this$ $currency$ $pair$ $in$ $the$ $second$ $table)$ in units of the currency being converted (i.e., here currency 1). Formulate the problem of finding the most cost effective method to convert these holdings into US$, as a min cost flow problem.

(2) Suppose there are transaction limits for converting yen, rupiah, and ringgits (only these currencies, no limits for converting other currencies) as shown in the following table (unit = equivalent of million US$). Then find the most cost effective way of converting as much of these currency holdings into US$ as possible.

<table>
<thead>
<tr>
<th>From</th>
<th>Transaction Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>JY</td>
<td>IR</td>
</tr>
<tr>
<td>JY</td>
<td>–</td>
</tr>
<tr>
<td>IR</td>
<td>5</td>
</tr>
<tr>
<td>MR</td>
<td>3</td>
</tr>
</tbody>
</table>

2.37. Blending of residual fuel oil: Crude oil is separated into many different products such as naphtha (used for making gasoline), diesel oil, etc. The most expensive products produced from crude oil are these higher volatile fractions. After all these highly profitable fractions are removed, one of the fractions that remains at the bottom is heavy residual fuel oil; its market price is quite low compared to the above products. The sulfur content in heavy fuel oil is usually high, and burning it requires special equipment; therefore, this fraction is mostly used as fuel in ships (i.e., ocean-going vessels) under the name “bunker fuel.” The quality of heavy fuel oil depends on the type of crude oil from which it is produced, and measured by three important characteristics: SG (specific gravity = weight/volume), SC (sulfur content, percent by weight), VBI (viscosity blend index, which blends linearly when
quantities are measured in volume units like gallons). To meet the specifications on these characteristics, bunker fuel sold to customers is usually a mixture of various fractions.

Suppose we have three different fractions available to blend into bunker fuel: AF (asphalt flux), CO (clarified oil), and K (kerosene). Data on them is given below.

<table>
<thead>
<tr>
<th>Fraction</th>
<th>Value of characteristic</th>
<th>Price $/gallon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SG</td>
<td>SC</td>
</tr>
<tr>
<td>AF</td>
<td>0.98</td>
<td>2.39</td>
</tr>
<tr>
<td>CO</td>
<td>0.91</td>
<td>2.20</td>
</tr>
<tr>
<td>K</td>
<td>1.20</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Bunker fuel can be sold at $1.80/gallon. Select appropriate units for measuring the three constituent fractions, and assuming that the blending cost is negligible, formulate an LP model for determining an optimum blend for bunker fuel.

2.38. Gasoline blending: A gasoline blending company has six types of raw gasolines available with data given below.

<table>
<thead>
<tr>
<th>Type</th>
<th>Oc. R. (octane rating)</th>
<th>Available b/day</th>
<th>Cost $/gallon</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>72</td>
<td>2,500</td>
<td>2.40</td>
</tr>
<tr>
<td>T2</td>
<td>81</td>
<td>3,800</td>
<td>2.55</td>
</tr>
<tr>
<td>T3</td>
<td>86</td>
<td>4,300</td>
<td>2.75</td>
</tr>
<tr>
<td>T4</td>
<td>89</td>
<td>5,500</td>
<td>2.85</td>
</tr>
<tr>
<td>T5</td>
<td>96</td>
<td>1,500</td>
<td>3.10</td>
</tr>
<tr>
<td>T6</td>
<td>99</td>
<td>1,000</td>
<td>3.20</td>
</tr>
</tbody>
</table>

1 barrel = 42 gallons.

They blend these raw gasolines into two grades of fuel: regular and premium. Minimum specifications on Oc. R. for regular is 85 and for premium it is 95. They need to produce a minimum of 9,000 b of regular/day and between 2,000 and 4,000 b/day of premium. The selling prices of these grades are $3.05/g and $3.20/g for regular and premium, respectively. Formulate an LP to determine an optimum blending plan for this company (Dantzing and Thappa (1997) of Chap. 1).

2.39. A production planning problem: Over the next 5-week horizon, a company has to produce and deliver to its customer a special commodity according to the following schedule: 200 units each at the end of weeks 1 and 2; then 300, 400, 500 units, respectively, at the end of weeks 3, 4, and 5.

The production process requires workers who received special training. The training of a new worker, carried out by a trained worker, takes 1 week. They begin week 1 with 10 trained workers on hand. Each trained worker can train up to five new workers during a week if assigned to training during that week, or produce up
to 10 units of the product if assigned to production during that week. Only workers trained at the company are used for production or training of new workers.

Shortages in delivery cost $100/unit short/week short, until delivered. By the end of the fifth week all required deliveries must be completed. A unit of product produced ahead of schedule must be stored at the company until its delivery time, this storage cost is $20/unit/week.

Salaries of workers are $15,000/week if used for production, $20,000/week if used to train new workers. Treating the number of workers used for the various tasks each week as continuous variables, formulate the problem of determining an optimum production/training schedule as an LP model (Dantzig (1963) of Chap. 1).

2.40. Minimizing the cost of production level changes: A company has a process that produces a product P. The production rate of this process measured in tons/month can be changed, but this change costs money.

The company’s policy is to change the production rate, if necessary, only at the beginning of each month. During any month, the process keeps on producing at the same rate as it was set at the beginning of that month.

To increase the production rate from existing level costs $c_1 = 100 per ton/month increase. To decrease production rate from existing level costs $c_2 = 50 per ton/month decrease. When the production rate is changed, the amount of increase or decrease in the rate has to be ≤ 100 tons.

The first month’s production rate can be set at any nonnegative level; there is no cost for this.

Production in a month can either be used for meeting the demand in that month or put in storage for meeting the demand later on. The storage room has a capacity for storing up to $s = 1,000$ tons and collects storage charges at the end of each month on the total amount in storage then, at the rate of $c_3 = 10/ton$. They have 100 tons in storage at the beginning of month 1.

Demand data ($d_i =$ demand (in tons) in period $i$: $i = 1$ to $6$) is (100, 500, 800, 600, 300, 500), and demand in each month has to be met exactly. At the end of month 6 they would like to have at least 250 tons in storage.

The actual production cost/ton is the same in all the months so the company needs to determine the production, storage plan in months 1–6 to minimize the sum of storage and production level changing costs while satisfying all the constraints. Formulate this as an LP and justify your formulation carefully.

2.41. $x = (x_1, \ldots, x_n)^T \in R^n$, where $n \geq 4$ is the vector of decision variables in the optimization problem

Minimize $\theta(x)$

subject to $Ax = b$

and $x \geq 0$,

where $A$ is a given $m \times n$ matrix and $\theta(x) = \text{Maximum}\{x_1, 5x_2, 10x_3\} - \text{Minimum}\{x_1, 5x_2, 10x_3\}$. 
Is it possible to transform this problem into a linear program? Why? If it is, give that transformation and a clear explanation why it will work.

2.42. Six period production planning problem: A company divides the year into six planning periods for planning purposes. Here is the relevant data for one important product of this company for the coming year.

<table>
<thead>
<tr>
<th>Period</th>
<th>Total Prod. cost ($/ton) = c_j</th>
<th>Prod. capacity (tons) k_j</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>15,000</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>20,000</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>22,000</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>30,000</td>
</tr>
<tr>
<td>5</td>
<td>35</td>
<td>27,000</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>25,000</td>
</tr>
</tbody>
</table>

There is unlimited demand for this product. The company has two main customers for it who have very different purchasing strategies.

Customer 1 wants to sign a purchasing deal for the whole year with the company, for any amount of the product, at $p_1 = 85$; but he stipulates a phased delivery: If $q_1$ tons is the quantity the company agrees to sell him, he wants 10% of this quantity delivered at the end of each of periods 1, 2, 5, 6, and 30% of this quantity delivered at the end of each of periods 3 and 4.

Customer 2 will purchase any quantity of the product available at the end of each period, but only at the rate $p_2 = 65$.

Product manufactured during a period has to be sold by the end of that period or disposed off, and it cannot be stored into the next period.

Formulate the problem of maximizing the net profit of the company during the year as an LP.

2.43. Speciality chemicals manufacturing: A speciality chemicals company makes three chemicals denoted by $C_1$, $C_2$, $C_3$: each in a separate division. These are rare chemicals used for highly specialized applications, every one who uses them always use all three of them in combination in proportions: $C_1:C_2:C_3 = 1:2:4$ (i.e., each unit of $C_1$ is always used in combination with 2 units of $C_2$ and 4 units of $C_3$, etc.). Also as they are very expensive chemicals, whenever a customer buys any of them, they always buy all of them in measured quantities in the above proportion.

Hence the customer demand for these chemicals can be measured unambiguously by the demand for any one of them (since the demand for the other two can be obtained from that one using the above proportions). Four expensive raw materials $RM_1$ to $RM_4$ are used in making these chemicals. Here is the relevant data for a period.
<table>
<thead>
<tr>
<th>Raw material</th>
<th>Units of input to make 1 unit</th>
<th>Max units available</th>
</tr>
</thead>
<tbody>
<tr>
<td>RM₁</td>
<td>2 0 1</td>
<td>3,000</td>
</tr>
<tr>
<td>RM₂</td>
<td>4 3 0</td>
<td>7,000</td>
</tr>
<tr>
<td>RM₃</td>
<td>0 3 1</td>
<td>5,000</td>
</tr>
<tr>
<td>RM₄</td>
<td>0 1 4</td>
<td>6,000</td>
</tr>
</tbody>
</table>

Labor man hours. 3 2 4

They can allocate up to 10,000 man hours of labor during regular time for the production of these chemicals, and up to another 3,000 man hours during overtime.

(1) Select suitable units for measuring customer demand for these chemicals, and formulate an LP model for determining the maximum customer demand that the company can meet in this period.

(2) The prices of raw materials RM₁ to RM₄ in money units/unit are 250, 175, 300, 400. The labor needed is highly skilled labor, regular time cost/man hour is 5, and overtime cost is 7.5 money units.

Assuming that the maximum customer demand that can be met (in units selected in (1)) is \( d \), formulate an LP model for meeting this demand exactly, at minimum cost.

2.44. Milk procurement: Organic Valley (OV) is a company that buys organic raw milk from individual farms in the milk-producing area of Wisconsin, divided into Regions 1 and 2. The farms vary in size, but all are small to medium. The composition of raw milk in terms of butter content and its “separation properties” tend to be more or less the same among the farms in each region, but the price/liter of raw milk varies from farm to farm. So, OV has classified the farms in each region into two classes, depending on the average price/liter of raw milk. Here is the data:

The average butter content of raw milk from Regions 1 and 2 is 40, 35 g/liter, respectively. The two classes in Region 1 are 1.1 and 1.2 and those in Region 2 are 2.1 and 2.2. In class 1.1 and 1.2, the average price (cents)/liter is 45 and 50, and the maximum availability in liters/day is 2,000, unlimited, respectively. In class 2.1 and 2.2, the average price (cents)/liter is 40 and 47, and the maximum availability in liters/day is 3,000 and unlimited, respectively.

In procuring milk each day, depending on the quantity to be purchased from each region, OV allocates that quantity among the two classes in that region so as to minimize the total cost.

For each region \( i = 1, 2 \), give the expression \( f_i(x_i) \) for the price of buying \( x_i \) liters of raw milk from region \( i \), explaining the reason for it.

Raw milk may be used in the blending process as it is or it may be passed through a separator before blending. Each liter from Region 1 when separated yields 0.5 L with butter 20 g/L, and another 0.5 L with butter 60 g/L at a cost of 3 cents/L. Similarly each liter from Region 2 when separated yields 0.5 L with butter 15 g/L, and another 0.5 L with butter 55 g/L at a cost of 3 cents/L.
The blending process blends the various milks into two products (grades of milk called cream and regular) with the following specifications.

<table>
<thead>
<tr>
<th>Product</th>
<th>Min. required butter content (g L⁻¹)</th>
<th>Mkt. price (cents L⁻¹)</th>
<th>Demand (L day⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cream</td>
<td>52</td>
<td>150</td>
<td>1,000</td>
</tr>
<tr>
<td>Regular</td>
<td>22</td>
<td>80</td>
<td>8,000</td>
</tr>
</tbody>
</table>

Any leftover unsold milk is sold to an ice cream plant for 38 cents/L as long as the butter content in it is ≥ 15 g/L.

Formulate the problem of maximizing OV’s net daily profit as an LP.

2.45. Milk blending: We have eight different grades of milk available with different butter contents. For \( i = 1 \) to 8, the butter content of grade \( i \) is 15, 20, 40, 45, 47, 50, 55, 60 g/L, respectively, and \( x_i \) liters is the quantity of this grade of milk available.

We have a customer to whom milk can be sold, but he will only buy milk in which the butter content is ≥ 42 g/L.

We have a blender that can blend any combination of several grades of milk. From the grades of milk that we have, it is required to determine the maximum quantity of milk satisfying the customer’s specification on butter content that we can produce using this blender. Formulate this problem as an LP.

2.46. Finding an interior point of a convex polyhedron: Let \( K = \{ x : Ax \geq b \} \). An \( x \in K \) is called an interior point of \( K \) if it satisfies \( Ax > b \), boundary point of \( K \) otherwise. Formulate the problem of finding an interior point of \( K \), or establishing that \( K \) has no interior point, as an LP.

On the other hand, if the goal is to find a boundary point of \( K \), how can one find it?

2.47. Optimizing bets at a horse race: A person has \( \beta \) to bet on \( n \) horses competing in a race. The race course has the following pay-off policy: If \( x_i \) $ is bet on the \( i \)th horse, the pay-off from this bet is \( \alpha_{0i} x_i \) if that horse comes first in the race, \( \alpha_{1i} x_i \) if it comes second in that race, 0 otherwise.

Formulate the problem of determining how much to bet on each horse in the race to maximize the minimum net gain, irrespective of which of the \( n \) horses come in first or second in the race.

2.48. Optimal revision of estimates to satisfy monotonicity: A sequence of parameters \( (a_1, \ldots, a_5) \), which are known to be monotonic increasing, is being estimated from data subject to random fluctuations. The estimation process estimates each \( a_i \) independently from the others. The estimates obtained are denoted by \( (\bar{a}_1, \ldots, \bar{a}_5) \), because of random fluctuations this sequence may not be monotonic increasing with \( i \). It is required to revise the estimate \( \bar{a}_i \) into \( x_i \), for \( i = 1 \) to 5, making the smallest possible changes, so that the revised sequence \( (x_1, \ldots, x_5) \) is monotonic increasing. Formulate this as an LP.
2.49. **Revising property tax rates:** A city currently collects property taxes within its boundary; this is the major source of its revenue. Having incurred budget deficits for several years, the city has appointed a committee to revise the property tax rates, and also suggest other taxes that the city can levy to balance the budget.

Besides property taxes (PT), the committee is investigating the possibility of including food and drug taxes (FDT), general sales tax (ST), and a gasoline tax (GT).

The tax base of property values is $550 million, and PT rate is an annual percentage of this value. The annual food and drug sales volume and general sales volume are $35 million and $55 million, respectively; again, the FDT and ST rates are percentages of these volumes. Annual gasoline sales volume within the city is estimated at 8 million gallons; the GT rate will be in the form of cents/gallon sale.

The committee wants to determine the PT, FDT, ST, GT rates so as to achieve the following goals as far as possible:

- **G1:** Total tax revenue should be $\geq 16$ million
- **G2:** FDT should be $\leq 10\%$ of total tax revenue
- **G3:** ST should be equal to $20\%$ of all taxes collected
- **G4:** GT should be $\leq 3$ cents/gallon
- **G5:** GT should be $\leq 12\%$ of the total tax revenue.

The committee has decided that missing **G1** (per million $\$, **G4** (per 1 cent/gallon), and **G2**, **G3**, or **G5** (per 1%) should all carry the same penalty.

Formulate the problem of deciding the PT, FDT, ST, GT rates so as to minimize the total penalty as an LP model.

2.50. **A property of convex functions:** Let $f(x)$ be a real-valued function in $x \in \mathbb{R}^n$.

Prove that if $f(x)$ is a convex function and $\alpha$ is any real number, then $\{x : f(x) \leq \alpha\}$ is the set of all points $x$ satisfying $f(x) \leq \alpha$ must be a convex set.

2.51. **Arranging financing for construction work:** The construction of a bridge starts on 1 Sept of year 1, and is expected to be completed on 31 Aug of year 5. For this, the city has agreed to make the following payments to the contractor.

<table>
<thead>
<tr>
<th>Date</th>
<th>Amount to be paid ($\text{million}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 September, year 1</td>
<td>12</td>
</tr>
<tr>
<td>1 September, year 2</td>
<td>14</td>
</tr>
<tr>
<td>1 September, year 3</td>
<td>18</td>
</tr>
<tr>
<td>1 September, year 4</td>
<td>15</td>
</tr>
</tbody>
</table>

The city will raise this money by selling bonds on each of the above dates. Let $x_j$ denote the $\text{Amount of bonds to be sold on the } j \text{th of these dates, } j = 1 \text{ to } 4$.

All bonds accumulate simple interest only on the principal amount. Four year bonds, three year bonds, two year bonds, one year bonds pay, respectively, 10\%, 9\%, 8\%, 7\% simple interest yearly. All bond amounts plus accrued interest will have to be paid to bondholders in a simple lump sum payment on 1 September of year 4.
If the money collected in a bond is not used to pay the contractor on the day the bond is sold, the city immediately invests it in a short-term (exactly one year) time deposit, which pays the principal +6% interest paid exactly one year later. So, on each of the dates in above tableau, contractors are paid out of money collected from bond sales on that day + any money coming from repayment of short-term deposit made 1 year before, and any remaining amount is invested in short term deposits right away.

The Federal government agreed to pay the city $59 million, the total construction cost of bridge, on 1 September of year 4. The city will use this money towards lump sum payments to be made to bondholders on that day.

It is required to determine optimum values of \( x_j \)'s and short term deposits to be made in each year of project horizon so as to minimize total interest money to be paid to bond holders on 1 September of year 4 when they get lump sum payments. Formulate this as an LP.

2.52. Managing a two-dam reservoir system: The state has two dams with reservoirs denoted by \( R_1, R_2 \), respectively, on a river. Water for the river comes mainly from melting snows on mountains before \( R_1 \). Estimated availability of water at \( R_1 \) is 294 MAF (million acrefeet)/year.

At \( R_1 \), the state has agreed to supply \( \geq 24 \) MAF/year to a town there. Remaining water at \( R_1 \) flows down to \( R_2 \) losing 20% to evaporation along the way. At \( R_2 \) state releases some water to farmers there for irrigation, and the rest passes through a hydroelectric generator and then flows further downstream.

The town at \( R_1 \) pays \$0.5/AF for water supplied. Farmers pay \$0.2/AF for irrigation water released to them. Water passing through hydroelectric generators earns \$0.8/AF.

The state has a target of 100 MAF/year for releasing to farmers for irrigation, and would be happy if this target can be exceeded. The penalty for downfall in irrigation water release below target is 10 penalty units/MAF downfall.

The state has a target of \$144 million for total income/year from this system, and would be happy if this target can be exceeded. Penalty for downfall in total income below target is 4 penalty units per \$million downfall.

Give a goal programming model for allocating available water to the three uses in this two-objective optimization problem.

2.53. \( f(x) \) is a convex function defined in the space of \( x = (x_1, x_2)^T \), satisfying following property:

**Property:** Inside the circle with \( x^* = (3, 2)^T \) as center and radius 1, \( x^* \) is the unique minimum point for \( f(x) \), that is, for all points \( \bar{x} \) inside the circle, \( \bar{x} \neq x^* \), we have \( f(\bar{x}) > f(x^*) \).

Take any point \( x \) outside this circle. Using Jensen’s inequality, prove that it is not possible for \( x \) to satisfy \( f(x) < f(x^*) \).

2.54. Managing Tunis Water Resources System: The Tunis Water Resources System (TWRS) comprises seven reservoirs interconnected by water conveyance
structures (WCS) (a canal and three pipelines). Each reservoir provides water for a number of demand centers, and a demand center may receive water from more than one reservoir.

The reservoirs (all with first initial “R”) are RJ (Reservoir Joumine, or Joumine), RBM (Ben Metir), RK (Kasseb), RBH (Bou Heurtma), RM (Mellegue), RSS (Sidi Salem), and RS (Siliana). Data on their capacities and the expected incremental inflows from precipitation in its catchment area in the month of May are given below:

<table>
<thead>
<tr>
<th>Reservoir</th>
<th>RJ</th>
<th>RBM</th>
<th>RK</th>
<th>RBH</th>
<th>RM</th>
<th>RSS</th>
<th>RS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity</td>
<td>120</td>
<td>44</td>
<td>72</td>
<td>103</td>
<td>89</td>
<td>510</td>
<td>62</td>
</tr>
<tr>
<td>(10^6 m^3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected May inflow (10^6 m^3)</td>
<td>2.9</td>
<td>0.9</td>
<td>1.7</td>
<td>3</td>
<td>15.3</td>
<td>19.6</td>
<td>2.5</td>
</tr>
</tbody>
</table>

There are demand centers all identified with “D” as the initial letter in their abbreviations, with demand for water in the month of May as given below.

<table>
<thead>
<tr>
<th>Demand center</th>
<th>Expected May demand (10^6 m^3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIMA</td>
<td>1.2</td>
</tr>
<tr>
<td>DBLI</td>
<td>0.9</td>
</tr>
<tr>
<td>DJE</td>
<td>2</td>
</tr>
<tr>
<td>DIBH</td>
<td>10.7</td>
</tr>
<tr>
<td>DINE</td>
<td>0.3</td>
</tr>
<tr>
<td>DIAEA</td>
<td>9.3</td>
</tr>
<tr>
<td>DISI</td>
<td>4.3</td>
</tr>
<tr>
<td>DIMSC</td>
<td>13.1</td>
</tr>
<tr>
<td>DIBV</td>
<td>1.8</td>
</tr>
<tr>
<td>DTO</td>
<td>2.6</td>
</tr>
<tr>
<td>DTU</td>
<td>5</td>
</tr>
</tbody>
</table>

The following Fig. 2.9 shows the WCS with its canals (in solid lines), pipelines (in dashed lines), with their monthly flow capacities in units of 10^6 m^3/month, and directions of flow indicated by arrows.

It is required to determine how much water should be supplied in the month of May to each demand center through the various canals and pipelines to meet all the demands, such that the amount of water in each reservoir at the end of May is as close as possible to the amount in it at the beginning of May. Formulate this as an LP. (Data from Nandalal and Bogardi 2007.)

2.55. Maximizing returns from short-term investments: In the automobile industry, companies make improvements in their car models and start selling the improved versions of the cars at the beginning of each year. These new versions
This figure shows the WCS with its canals (in solid lines), pipelines (in dashed lines), with their monthly flow capacities in units of $10^6$ m$^3$/month, and directions of flow indicated by arrows. Each draws water from a reservoir node, and as it passes through each demand node satisfying the demand there, the quantity of water it carries keeps decreasing. For the sake of this problem, we ignore evaporation losses.

enjoy high sale volumes during the first half of the year, after which the sale volumes tend to level off during the second half of the year. That is why income to these companies in the form of sale revenues from dealers tends to be higher in the first half of the year than during the second half.

By the middle of the year, quality problems with the versions of car models being sold become well known, and companies then set up new design teams to work on fixing up these problems in the new versions of these car models to be released in the market at the beginning of next year. That is why design changing expenses for these companies tend to be higher in the second half of the year than during the first half.

Hence, the income for these companies tends to be higher than their expenses during the first half of the year and vice versa during the second half. These companies normally operate on annual accounting, and their profits for the year are declared only at the end of the year. So, to maximize their annual profits, these companies normally invest their surplus income over expenses in the first half of the year, in short term investment opportunities, to be cashed as needed or by the end of the year.
We provide the data for a company that divides the year into four periods of three months each. Data on the expected income and expenses in each period is given below. There are two investment opportunities, $I_1$ and $I_2$ for short-term investments. In each of these opportunities, any amounts can be invested for periods ranging from one, two, or three periods (can assume that all transactions take place only on the last day of each period) with fractional returns given in a table below (i.e., if $x$ is the amount invested, $r$ is the fractional return for a duration of $k$ periods, and this investment will pay out $x(1 + r)$ at the end of $k$ periods after investing).

<table>
<thead>
<tr>
<th>Period</th>
<th>Expected income</th>
<th>Expected expenses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>200</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>180</td>
<td>110</td>
</tr>
<tr>
<td>3</td>
<td>150</td>
<td>180</td>
</tr>
<tr>
<td>4</td>
<td>170</td>
<td>250</td>
</tr>
<tr>
<td>Annual total</td>
<td>700</td>
<td>640</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Duration (no. periods)</th>
<th>Fractional returns in $I_1$</th>
<th>Fractional returns in $I_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>2</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>3</td>
<td>0.08</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Remember that all investments during the year will be cashed by the end of the year. Also, the maximum invested over the year in any investment opportunity cannot exceed 150 money units. The only money that the company invests in these investments is any money on hand (from income in that period either from dealers or from earlier investments that are cashed). Formulate the problem of determining how much the company can invest in each investment opportunity, so as to maximize the money left over in the companies hands by the end of the year.

**2.56. Cake mix blending:** A premium-quality cake mix is a blend of six different powdered ingredients called $I_1$ to $I_6$.

There are three weighing machines (WM-1 to WM-3) for measuring the quantities of the ingredients. Each weighing machine has one or more sizeable hoppers attached to it, each hopper holding one of the ingredients.

The following table indicates the ingredients contained in the hoppers attached to each of the weighing machines (* mark indicates that one hopper of the corresponding weighing machine contains the ingredient, blank entry indicates that none of the hoppers on this weighing machine contains that ingredient).

<table>
<thead>
<tr>
<th></th>
<th>$I_1$</th>
<th>$I_2$</th>
<th>$I_3$</th>
<th>$I_4$</th>
<th>$I_5$</th>
<th>$I_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WM-1</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WM-2</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WM-3</td>
<td>*</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>
An ingredient can be weighed on a weighing machine only if one of its hoppers holds that ingredient. For example, on WM-1 we can weigh ingredients $I_1$, $I_3$, $I_4$, but none of the others. Also, $I_1$ can be weighed on WM-1 or WM-3 or both of these, but not on WM-2.

On each weighing machine, the ingredients are emptied from one of its hoppers at a time, each hopper emptying for a preset duration depending on the quantity of that ingredient to be measured on that weighing machine into the scale of that weighing machine. After all the hoppers finish delivering their ingredients, the contents of the scale are conveyed to a bin where they will be blended into the batch.

The weighing machines work simultaneously and independently in the plant, and the whole operation is computer controlled. As soon as the scales on all the three weighing machines deliver their contents to the bin, the batch is complete, weighing stops, the scales are cleaned, and preparations take place for the next batch.

Some ingredients flow faster than the others from the hoppers.

$$g_i = \text{time in seconds for 1 kg of ingredient } I_i \text{ to flow from the hoppers onto the scales of a weighing machine, for } i = 1 \text{ to } 6.$$ 

Here is the data:

<table>
<thead>
<tr>
<th>Ingredient</th>
<th>$I_1$</th>
<th>$I_2$</th>
<th>$I_3$</th>
<th>$I_4$</th>
<th>$I_5$</th>
<th>$I_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>kg needed for a batch</td>
<td>130</td>
<td>140</td>
<td>101</td>
<td>290</td>
<td>29</td>
<td>21</td>
</tr>
<tr>
<td>$g_i$</td>
<td>1</td>
<td>1.2</td>
<td>1.3</td>
<td>1.1</td>
<td>0.8</td>
<td>0.9</td>
</tr>
</tbody>
</table>

The capacity of the scales on each of the three weighing machines is 400 kg.

The problem is to determine how much quantity of each ingredient should be weighed on each weighing machine in order to minimize the clock time for weighing a batch.

Give a formulation of this problem as a linear program. Define all your decision variables very clearly, and write the objective function and all the constraints carefully. Justify your model.

2.57. In $R^2$, the rectilinear distance between two points $a = (a_1,a_2)$ and $b = (b_1,b_2)$ is defined to be $d(a,b) = |a_1 - b_1| + |a_2 - b_2|$.

We are given four points in $R^2$. These are $a^1 = (-10, 20)$, $a^2 = (60, 30)$, $a^3 = (40, 10)$, $a^4 = (80, 60)$. The weights associated with these points are $(w_1, w_2, w_3, w_4) = (0.2, 0.1, 0.3, 0.4)$, respectively.

It is required to find a point $x \in R^2$ to minimize $z(x) = \max\{w_t d(x, a^t) : t = 1 \text{ to } 4 \}$. Give a linear programming formulation of this problem.

2.58. Operating a hydroelectric power system: An agency operates two water reservoirs on a river system. Each of these reservoirs has a hydroelectric power generator, and all the water released from this reservoir flows through this generator and generates power. AF = acrefeet, KAF = kiloacrefeet or 1,000 AF are units for measuring the volume of water. KWH = kilowatthour is a unit for measuring electric power.
The amount of power generated when an AF of water is released from the reservoir depends nonlinearly on the *head* which is determined by the amount of water in the reservoir at the time of release; however, to model the problem as a linear program, we will assume that it is a constant = the average amount given in the following table.

For operational decisions, a year is divided into six periods (or seasons) of two months each. For simplicity, assume that each period is an interval of time with uniform characteristics, that is, the inflows during this period from precipitation occur at a constant rate.

During each period, the agency is required to release an amount of water ≥ a specified minimum amount from each reservoir for downstream uses. All water released from reservoirs flows through the power generators and hence produces electricity. During each period that water is released at a constant rate so that the total release for the period equals the planned amount.

At any instant of time, if reservoir is at its full capacity, additional inflowing water spills over the dam. The spilled water does not generate any electric power. As an example, if water stored in Reservoir 1 with its capacity of 4,000 KAF at the beginning of a period is 3,200 KAF, the inflow during this period into this reservoir is 1,400 KAF, and the water released during this period from this reservoir is 200 KAF, then $3200 + 1400 - 200 - 4000 = 400 KAF$ will be the spilled water from this reservoir during this period.

Here is the data on the two reservoirs:

<table>
<thead>
<tr>
<th>Reservoir</th>
<th>Capacity</th>
<th>Power generated by releasing 1 AF&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Water stored at year beginning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4000 KAF</td>
<td>310 KWH</td>
<td>1800 KAF</td>
</tr>
<tr>
<td>2</td>
<td>6000 KAF</td>
<td>420 KWH</td>
<td>2500 KAF</td>
</tr>
</tbody>
</table>

<sup>a</sup>Average amount per AF as mentioned above

**Power sales:** They have an industrial customer (IC) who buys power on an annual basis and pays $70/1,000 KWH. The IC stipulates that if the annual KWH of power sold to them is denoted by $e_0$, then $p_i e_0$ KWH of it must be delivered in period $i$, for $i = 1$ to 6; where the $p_i$ are given in the following table. In each period, any power not sent to the IC is sold in that period itself to the regional power grid at a price of $50/1,000$ KWH.

<table>
<thead>
<tr>
<th>Period</th>
<th>Inflow (KAF) into reservoir</th>
<th>Min release, KAF from reservoir</th>
<th>$p_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i$</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>550</td>
<td>2620</td>
<td>200</td>
</tr>
<tr>
<td>2</td>
<td>1,470</td>
<td>2300</td>
<td>200</td>
</tr>
<tr>
<td>3</td>
<td>990</td>
<td>1230</td>
<td>200</td>
</tr>
<tr>
<td>4</td>
<td>150</td>
<td>730</td>
<td>200</td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>410</td>
<td>200</td>
</tr>
<tr>
<td>6</td>
<td>160</td>
<td>500</td>
<td>200</td>
</tr>
</tbody>
</table>
At the end of the year, the amount of water in storage in each reservoir must be the same as that at the beginning of the year.

It is required to determine how much water to release from each reservoir in each period, the annual amount of power to sell IC, and the amount of power to sell the regional power grid in each period, so as to maximize the total annual return from power sales subject to all the above constraints. Formulate this as an LP, but do not solve numerically.

2.59. Consider the LP: Minimize some objective function, $z = cx$ subject to the following constraints in standard form.

<table>
<thead>
<tr>
<th></th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$x_4$</th>
<th>$x_5$</th>
<th>$x_6$</th>
<th>$x_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>$-1$</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>$-2$</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>$-2$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>1</td>
<td>$-1$</td>
</tr>
<tr>
<td>$x_j \geq 0$ for all $j$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Assuming that the problem has an optimum solution, prove that there must be an optimum solution of the problem in which at least one of the two variables $x_6$ or $x_7$ is 0.

2.60. Consider a company making five different car models indexed by $i = 1$ to 5, in the luxury car segment. In the market there are 20 other car models indexed by $i = 6$ to 25, made by other manufacturers in this segment.

The sales volume $Q_i$ during a quarter (of the year) of any of these car models depends on the sale price $P_i$ of this model and the sale prices of other car models in this segment.

The sales volume $Q_i$ during a quarter (of the year) of any of these car models depends on the sale price $P_i$ of this model and the sale prices of other car models in this segment. For a particular quarter, we are given the following data:

For $i = 1$ to 5, $S_i$ (= unsold inventory in the dealer lots at the beginning of the quarter) and $D_0^i$ (projected demand at current price levels) of car model $i$

$C = \text{total production capacity of car models } i = 1 \text{ to } 5 \text{ at the company's plant in this quarter}$

For $i = 1$ to 5, $P_i^0 = \text{price of car model } i \text{ at the end of the previous quarter}$

For $i = 6$ to 25, $P_i = \text{best estimate of the price of car model } i \text{ of the competitor for this quarter}$,

$a_{ij}$ for $i = 1$ to 5 and $j = 1$ to 25 is the rate of change in the expected demand of car model $i$ in this quarter per unit change in $P_j$,

$c_i$ for $i = 1$ to 5 is the total production cost/unit of the $i$th car model.

Since these car models are competing with each other in the marketplace, change in the selling price of any one of them affects the demand for the other car models too in this quarter; the $a_{ij}$ coefficients measure these effects, and these coefficients are estimated from market surveys on customers buying new cars. So, what we are assuming is that the expected demand in the quarter for the $i$th car model for $i = 1$ to 5 is $D_i = D_0^i - \sum_{j=1}^{25} a_{ij} P_j$, where $P_6$ to $P_{25}$ are given data and $P_1$ to $P_5$ are the prices set for car models $i = 1$ to 5, respectively, for this quarter, these are
2.6 Exercises

Let $I_i$ denote the production level of the $i$th car model in this quarter by the company for $i = 1$ to 5. Then the amount $S_i + I_i$ is the volume available for sale of this model during this quarter. Assuming that the sales and production volumes of each car model are continuous variables, formulate the problem of determining $P_i, I_i$ to maximize the company’s net profit during the quarter (problem formulated by Robert Bordley).

2.61. $A$ is a matrix of order $m \times n$ of full row rank, and $S$ is the subspace of $R^n$, $S = \{x \in R^n : Ax = 0\}$.

For any pair of points $p = (p_j), q = (q_j) \in R^n$, the $L_1$ distance between them, denoted by $||p - q||_1$, is defined to be $\sum_{j=1}^n |p_j - q_j|$.

Given a point $y \in R^n$, $y \not\in S$, the $L_1$ projection of $y$ onto the subspace $S$ is defined as an optimum solution of the problem of minimizing $||y - x||_1$ over $x \in S$. Formulate the problem of finding an $L_1$ projection of a given point $y \in R^n$, $y \not\in S$ onto the subspace $S$ as a linear programming problem.

Is this $L_1$ projection of $y$ onto $S$ a unique point? Illustrate with a numerical example (from Brooks and Dula 2008).

2.62. Consider the problem discussed in Exercise 2.61. If $S$ is a hyperplane in $R^n$ containing the origin, that is, $S = \{x : a_1x_1 + \ldots + a_nx_n = 0\}$, where $(a_1, \ldots, a_n) \neq 0$ and $y \in R^n, y \not\in S$; derive the set of all points in $S$ which are at the minimum $L_1$-distance to $y$ (J.H. Dula).

2.63. Check whether the real-valued function $f(x_1, x_2) = -x_1 + (x_1 + x_2)e^{-(x_1 + x_2)}$ is a convex function (R. Saigal).

2.64. $w = (w_1, \ldots, w_n) > 0$ is a given vector of positive weights. For $y = (y_1, \ldots, y_n)^T$ let

$$L(y) = \text{Maximum}\{w_i y_i : \ i = 1, \ldots, n\}$$

$a \in R^n$ is given, and we are required to find $x = (x_1, \ldots, x_n)^T$ to minimize $L(a - x)$ subject to $x_j - x_{j+1} \leq 0$ for all $j = 1$ to $n - 1$. Formulate this problem as an LP.

How does the formulation change if we have additional constraints: $0 \leq x_j \leq 1$ for all $j$? (V.A. Ubhaya).

2.65. Let $\lambda \in R^1$ and $\epsilon > 0$ is a positive constant. Define $\theta(\lambda) = \epsilon \log\{1 + e^{[g(\lambda)/\epsilon]} + 1\}$, where $g(\lambda) = a + b\lambda$, an affine function. Show that $\theta(\lambda)$ is a convex function of $\lambda$. 

Decision variables the company can choose (so, for $i = 1$ to $5$, $P_i - P_i^0$ is the change in prices of these models from last quarter to this quarter through rebates or price increases).
References

Iverstine and Kinard (1977) Cases in production and operations management. Charles E. Meril Publishing Co., Columbus, Ohio
Lanzenauer CHV (1975) Cases in OR. Holden-Day, San Francisco
Lee SM (1972) Goal programming for decision analysis. Auerbach, Philadelphia
Sponk J (1981) Interactive multiple goal programming: Applications to financial management. Martinus Nijhoff, Boston
References


Optimization for Decision Making
Linear and Quadratic Models
Murty, K.G.
2010, XXVI, 482 p. 47 illus., Hardcover