IIT Hyderabad **G V Krishna Chaitanya (2021111023)**
- **S Sai Satvik (2021101117)**

Optimization Methods(CS1.404) **Lecture 9**
Instructor: Dr. Naresh Manwani 6^{th} **Feb 2025**

Outline. *This lecture focuses on the proofs surrounding the Supporting Hyperplane Theorem, Convex Hull, and Convex Cone. It also delves into Farkas' Lemma and a brief introduction to Convex Functions.*

1 Supporting Hyperplane Theorem:

Suppose $C \subset \mathbb{R}^n$, and $\bar{x}_0$ is a point in its boundary bd(C), i.e.,

$$\bar{x}_0 \in \text{bd}(C) = \text{cl}(C) \setminus \text{int}(C).$$

If $\bar{a} \neq 0$ satisfies

$$\bar{a}^T \bar{x} \leq \bar{a}^T \bar{x}_0, \quad \forall \bar{x} \in C,$$

then the hyperplane

$$\{\bar{x} \mid \bar{a}^T \bar{x} = \bar{a}^T \bar{x}_0\}$$

is called a *supporting hyperplane* to C at the point $\bar{x}_0$. This is equivalent to saying that the point $\bar{x}_0$ and the set C are separated by the hyperplane $\{\bar{x} \mid \bar{a}^T \bar{x} = \bar{a}^T \bar{x}_0\}$.
The geometric interpretation is that the hyperplane $\{\bar{x} \mid \bar{a}^T \bar{x} = \bar{a}^T \bar{x}_0\}$ is tangent to C at $\bar{x}_0$, and the halfspace $\{\bar{x} \mid \bar{a}^T \bar{x} \leq \bar{a}^T \bar{x}_0\}$ contains C.
A basic result, called the **supporting hyperplane theorem**, states that for any nonempty convex set C, and any $\bar{x}_0 \in \text{bd}(C)$, there exists a supporting hyperplane to C at $\bar{x}_0$. The supporting hyperplane theorem is readily proved from the separating hyperplane theorem.

1.1 Proof:

Consider a sequence of points $\bar{y}_0, \bar{y}_1, \bar{y}_2, \dots$ that do not belong to the interior of C and satisfy:

$$\lim_{k \to \infty} \bar{y}_k = \bar{x}_0,$$

where $\bar{x}_0$ is a boundary point of C.
For each $\bar{y}_k$, there exists a closest point $\bar{z}_k$ in C from $\bar{y}_k$ by the closest point theorem:

$$\bar{z}_k = \arg\min_{\bar{x} \in C} \|\bar{x} - \bar{y}_k\|.$$

By the convexity of C, we obtain:

$$\langle \bar{x} - \bar{z}_k, \bar{y}_k - \bar{z}_k \rangle \leq 0, \quad \forall \bar{x} \in C.$$

Define:

$$\bar{a}_k = \frac{\bar{y}_k - \bar{z}_k}{\|\bar{y}_k - \bar{z}_k\|}.$$

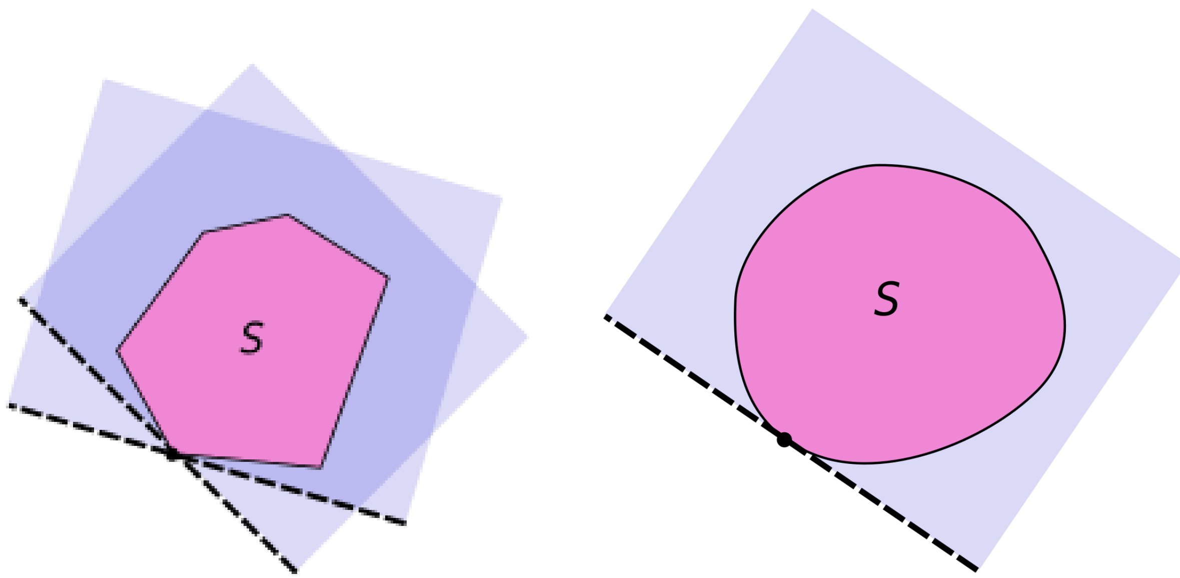

Figure 1: Some graphical examples of Supporting Hyperplanes

By the convexity of C, we obtain:

$$\langle \bar{x} - \bar{z}_k, \bar{a}_k \rangle \leq 0, \quad \forall \bar{x} \in C.$$

Since the sequence $\{\bar{a}_k\}$ is bounded, by the Bolzano-Weierstrass theorem, there exists a convergent subsequence $\{\bar{a}_i\}_{i \in I}$ such that:

$$\bar{a}_i \to \bar{a}^*.$$

Taking the limit as $i \to \infty$, we obtain:

$$\langle \bar{x} - \bar{z}_i, \bar{a}_i \rangle \leq 0.$$

Expanding,

$$\langle \bar{z}_i, \bar{a}_i \rangle = \langle \bar{z}_i - \bar{y}_i, \bar{a}_i \rangle + \langle \bar{y}_i, \bar{a}_i \rangle = -\|\bar{y}_i - \bar{z}_i\| + \langle \bar{y}_i, \bar{a}_i \rangle.$$

Thus,

$$\langle \bar{z}_i, \bar{a}_i \rangle < \langle \bar{y}_i, \bar{a}_i \rangle.$$

Since $\bar{a}_i(\bar{x} - \bar{z}_i) \leq 0$, we get:

$$\langle \bar{x}, \bar{a}_i \rangle \leq \langle \bar{z}_i, \bar{a}_i \rangle < \langle \bar{y}_i, \bar{a}_i \rangle.$$

Taking the limit,

$$\lim \langle \bar{x}, \bar{a}_i \rangle < \lim \langle \bar{y}_i, \bar{a}_i \rangle.$$

Thus,

$$\langle \bar{x}, \bar{a}^* \rangle \leq \langle \bar{x}_0, \bar{a}^* \rangle.$$

2 Convex Hull

Let X be a subset of $\mathbb{R}^n$. The *convex hull* of X, denoted as $\text{conv}(X)$, is the intersection of all convex sets containing X. That is,

$$\text{conv}(X) = \bigcap \{C \subset \mathbb{R}^n \mid C \text{ is convex and } X \subseteq C\}.$$

Lemma: The convex hull of a set of vectors $\{\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m\}$ is the set of all convex combinations of these vectors. That is,

$$\text{conv}(\{\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m\}) = \left\{ \sum_{i=1}^{m} \lambda_i \bar{x}_i \mid \lambda_i \geq 0, \sum_{i=1}^{m} \lambda_i = 1 \right\}.$$

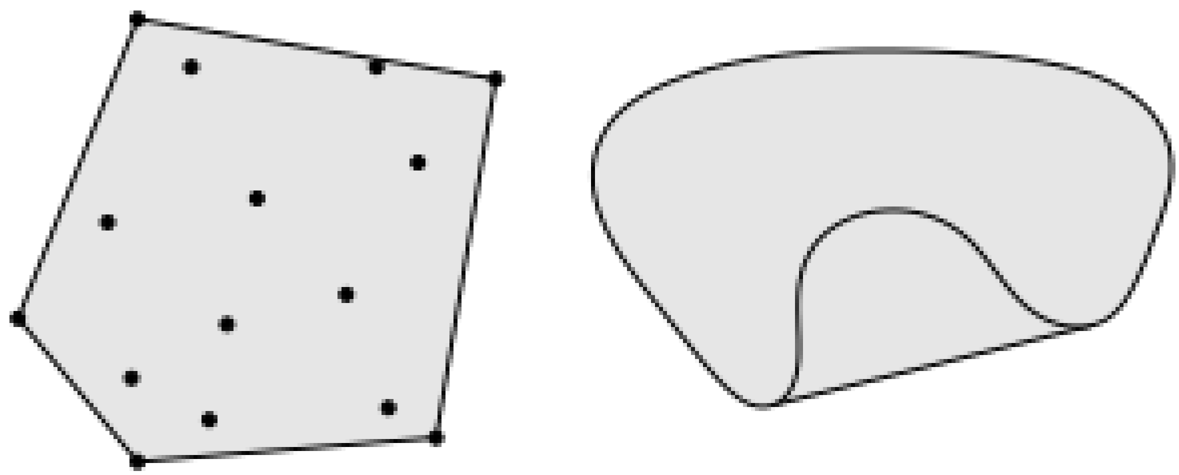

Figure 2: Some graphical examples of Convex Hulls

2.1 Proof that the Convex Hull is Convex

The convex hull of X is, by definition, an intersection of convex sets.
Let $X_1, X_2, \ldots, X_k$ be all convex sets that contain X. Let $\text{conv}(X)$ denote the convex hull of X.
Let $\bar{z}_1, \bar{z}_2 \in \text{conv}(X)$. Since $\text{conv}(X)$ is the intersection of all convex sets containing X, we have $\bar{z}_1, \bar{z}_2 \in X_i$ for all $i \in \{1, \ldots, k\}$.
As each X_i is convex, it follows that for any $\lambda \in [0, 1]$,

$$\lambda \bar{z}_1 + (1 - \lambda) \bar{z}_2 \in X_i, \quad \forall i \in \{1, \ldots, k\}.$$

Since this holds for all i, we conclude that

$$\lambda \bar{z}_1 + (1 - \lambda) \bar{z}_2 \in \bigcap_{i=1}^{k} X_i = \text{conv}(X), \quad \forall \lambda \in [0, 1].$$

Thus, $\text{conv}(X)$ is convex. $\square$

2.2 Proof of Lemma

We prove this by induction on m, the number of vectors in the set.
Base Case ($m = 1$): If $X = \{\bar{x}_1\}$, then the only possible convex combination is $\lambda_1 \bar{x}_1$ with $\lambda_1 = 1$. Thus, $\text{conv}(X) = \{\bar{x}_1\}$, which is trivially the set of all convex combinations of $\bar{x}_1$.
Inductive Step: Assume the lemma holds for any set of m vectors, i.e.,

$$\text{conv}(\{\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m\}) = \left\{ \sum_{i=1}^{m} \lambda_i \bar{x}_i \mid \lambda_i \geq 0, \sum_{i=1}^{m} \lambda_i = 1 \right\}.$$

We need to show that it holds for $m+1$ vectors. Consider the set $\{\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m, \bar{x}_{m+1}\}$. Any convex combination of these vectors can be rewritten as:

$$\bar{y} = \sum_{i=1}^{m+1} \lambda_i \bar{x}_i, \quad \text{where } \lambda_i \geq 0 \text{ and } \sum_{i=1}^{m+1} \lambda_i = 1.$$

Define $\mu = \sum_{i=1}^{m} \lambda_i$, which satisfies $0 \leq \mu \leq 1$. Then we can rewrite $\bar{y}$ as:

$$\bar{y} = \mu \sum_{i=1}^{m} \frac{\lambda_i}{\mu} \bar{x}_i + (1 - \mu) \bar{x}_{m+1},$$

where the terms inside the summation form a convex combination of $\bar{x}_1, \ldots, \bar{x}_m$. By the induction hypothesis, this belongs to $\text{conv}(\{\bar{x}_1, \ldots, \bar{x}_m\})$. Thus, $\bar{y}$ is a convex combination of an element of $\text{conv}(\{\bar{x}_1, \ldots, \bar{x}_m\})$ and $\bar{x}_{m+1}$, meaning it belongs to $\text{conv}(\{\bar{x}_1, \ldots, \bar{x}_m, \bar{x}_{m+1}\})$ i.e.

$$\left\{ \sum_{i=1}^{m+1} \lambda_i \bar{x}_i \mid \lambda_i \geq 0, \sum_{i=1}^{m+1} \lambda_i = 1 \right\} \subseteq \text{conv}(\{\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{m+1}\}) \tag{1}$$

The convex hull is the intersection of all convex sets that contain X. This implies that it belongs to the set of convex combinations of points in X:

$$\text{conv}(\{\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{m+1}\}) \subseteq \left\{ \sum_{i=1}^{m+1} \lambda_i \bar{x}_i \mid \lambda_i \geq 0, \sum_{i=1}^{m+1} \lambda_i = 1 \right\}. \tag{2}$$

From (1) and (2), we determine that

$$\text{conv}(\{\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{m+1}\}) = \left\{ \sum_{i=1}^{m+1} \lambda_i \bar{x}_i \mid \lambda_i \geq 0, \sum_{i=1}^{m+1} \lambda_i = 1 \right\}.$$

Thus, the lemma holds for $m+1$, completing the proof by induction.

3 Convex Cone

A set C is called a *cone* (or *nonnegative homogeneous*) if for every $\bar{x} \in C$ and $\theta \geq 0$, we have $\theta \bar{x} \in C$.
A set C is a *convex cone* if it is both convex and a cone, meaning that for any $\bar{x}_1, \bar{x}_2 \in C$ and $\theta_1, \theta_2 \geq 0$, we have

$$\theta_1 \bar{x}_1 + \theta_2 \bar{x}_2 \in C.$$

Geometrically, points of this form in two dimensions create a *pie slice* with its apex at the origin and edges passing through $\bar{x}_1$ and $\bar{x}_2$ (see Figure 2.4).
A point of the form $\theta_1\bar{x}_1 + \cdots + \theta_k\bar{x}_k$ with $\theta_i \geq 0$ is called a *conic combination* (or a *nonnegative linear combination*) of $\bar{x}_1, \ldots, \bar{x}_k$.
If $\bar{x}_i$ are in a convex cone C, then every conic combination of $\bar{x}_i$ is also in C. Conversely, a set C is a convex cone if and only if it contains all conic combinations of its elements. Similar to convex (or affine) combinations, the idea of conic combination extends to infinite sums and integrals.
The *conic hull* of a set C is the set of all conic combinations of points in C, i.e.,

$$\mathrm{cone}(C) = \left\{ \sum_{i=1}^{k} \theta_i \bar{x}_i \mid \bar{x}_i \in C, \theta_i \geq 0, i = 1, \ldots, k \right\},$$

which is also the smallest convex cone that contains C.

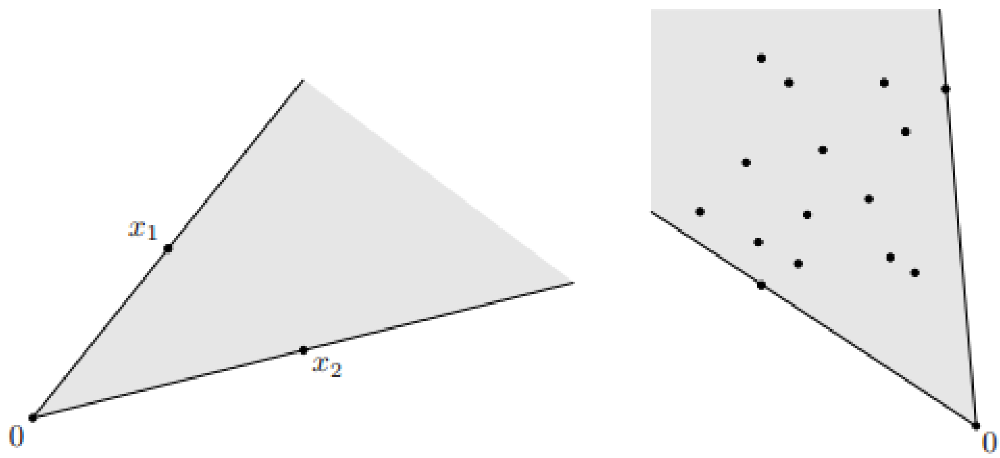

Figure 3: Some graphical examples of Convex Cones

4 Farkas' Lemma

Let $A \in \mathbb{R}^{m \times n}$ be a given matrix, and define the convex cone

$$C = \{A\bar{x} \mid \bar{x} \geq 0\}.$$

That is, C consists of all non-negative linear combinations of the columns of A, making it a closed convex cone.
Now, let $\bar{b} \in \mathbb{R}^m$ and $\bar{y} \in \mathbb{R}^m$ be arbitrary vectors. Then, exactly one of the following two mutually exclusive systems has a solution:

1. **Primal Feasibility Condition**: There exists $\bar{x} \in \mathbb{R}^n$ such that

$$A\bar{x} = \bar{b}, \quad \bar{x} \geq 0.$$

   This holds when $\bar{b}$ belongs to the convex cone generated by the columns of A, meaning that $\bar{b}$ can be expressed as a non-negative linear combination of the columns of A.

2. **Dual Feasibility Condition (Farkas' Alternative)**: There exists $\bar{y} \in \mathbb{R}^m$ such that

$$A^T \bar{y} \geq 0, \quad \bar{b}^T \bar{y} < 0.$$

This occurs when $\bar{b}$ does not belong to the convex cone of the columns of A, meaning that there exists a separating hyperplane (determined by $\bar{y}$) that strictly separates $\bar{b}$ from the cone.

4.1 Explanation and Intuition

The intuition behind this lemma is related to separating hyperplanes and duality principles. If the equation $A\bar{x} = \bar{b}$ with $\bar{x} \geq 0$ has no solution, then there exists a hyperplane (represented by $A^T \bar{y}$) that separates $\bar{b}$ from the cone generated by A. This separating hyperplane acts as a certificate proving the infeasibility of the first condition.

4.2 Proof

Assume that the first equation, $A\bar{x} = \bar{b}$ with $\bar{x} \geq 0$, has a solution. This means that $\bar{b}$ lies in the convex cone formed by the columns of A. So, there exists $\bar{x}_0 \geq 0$ such that $A\bar{x}_0 = \bar{b}$.
Now, consider any $\bar{y} \in \mathbb{R}^m$ such that $A^T \bar{y} \geq 0$. Taking the inner product with $\bar{b}$, we get:

$$\langle \bar{b}, \bar{y} \rangle = \langle A\bar{x}_0, \bar{y} \rangle = \langle \bar{x}_0, A^T \bar{y} \rangle \geq 0$$

which implies that the second system cannot hold.
Conversely, suppose the first system has no solution. Define the set

$$C = \{\bar{z} \in \mathbb{R}^m \mid \bar{z} = A\bar{x}, \bar{x} \geq 0\}.$$

Since $\bar{b} \notin C$, by the Separating Hyperplane Theorem, there exists a hyperplane that separates $\bar{b}$ from C. That is, there exists a vector $\bar{a} \in \mathbb{R}^m$ and a scalar p such that:

$$\langle \bar{a}, \bar{z} \rangle \geq p \quad \forall \bar{z} \in C, \quad \text{and} \quad \langle \bar{a}, \bar{b} \rangle < p.$$

Since $0 \in C$, substituting $\bar{z} = 0$ gives $\bar{a}^T 0 \geq p$, which implies $p \leq 0$. Combining this with $\bar{a}^T \bar{b} < p$, we get:

$$\langle \bar{a}, \bar{b} \rangle < 0.$$

Moreover, for any $\bar{x} \geq 0$, we have $A\bar{x} \in C$ and thus

$$\langle \bar{a}, A\bar{x} \rangle \geq p.$$

Since $p \leq 0$ and $\bar{x} \geq 0$, this implies that $A^T \bar{a} \geq 0$.
Thus, the hyperplane satisfies the conditions $\bar{a}^T \bar{b} < 0$ and $A^T \bar{a} \geq 0$, proving the second condition of the lemma.

5 Convex Functions

A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if $\operatorname{dom} f$ is a convex set and if for all $\bar{x}, \bar{y} \in \operatorname{dom} f$ and θ with $0 \le \theta \le 1$, we have

$$f(\theta\bar{x} + (1-\theta)\bar{y}) \le \theta f(\bar{x}) + (1-\theta) f(\bar{y}).$$

Geometrically, this inequality means that the line segment between $(\bar{x}, f(\bar{x}))$ and $(\bar{y}, f(\bar{y}))$, which is the chord from $\bar{x}$ to $\bar{y}$, lies above the graph of f.
A function f is strictly convex if strict inequality holds in the above equation whenever $\bar{x} \neq \bar{y}$ and $0 < \theta < 1$. We say f is concave if $-f$ is convex, and strictly concave if $-f$ is strictly convex.
For an affine function, we always have equality in the convexity condition, so all affine (and therefore also linear) functions are both convex and concave. Conversely, any function that is both convex and concave is affine.
A function is convex if and only if it is convex when restricted to any line that intersects its domain. In other words, f is convex if and only if for all $\bar{x} \in \operatorname{dom} f$ and all $\bar{v}$, the function $g(t) = f(\bar{x} + t\bar{v})$ is convex on its domain, $\{t \mid \bar{x} + t\bar{v} \in \operatorname{dom} f\}$.
This property is very useful since it allows us to check whether a function is convex by restricting it to a line.
The analysis of convex functions is a well-developed field. One simple result, for example, is that a convex function is continuous on the relative interior of its domain; it can have discontinuities only on its relative boundary.

5.1 Examples:

1) Norm is a Convex Function

Let $f(\bar{x}) = \|\bar{x}\|$. We show that $f(\bar{x})$ satisfies the definition of convexity.

$$f(\lambda\bar{x} + (1-\lambda)\bar{y}) = \|\lambda\bar{x} + (1-\lambda)\bar{y}\|$$

Using the triangle inequality:

$$\|\lambda\bar{x} + (1-\lambda)\bar{y}\| \le \|\lambda\bar{x}\| + \|(1-\lambda)\bar{y}\|$$

Since norms are positively homogeneous:

$$= \lambda\|\bar{x}\| + (1-\lambda)\|\bar{y}\|$$

$$= \lambda f(\bar{x}) + (1-\lambda) f(\bar{y})$$

Since this holds for all $\bar{x}, \bar{y}$ and $\lambda \in [0,1]$, we conclude that the norm function is convex.

2) Is $\bar{x}_1\bar{x}_2$ convex in the positive quadrant?

Let $f(\bar{x}) = \bar{x}_1\bar{x}_2$. We examine whether $f(\bar{x})$ is convex over

$$\Omega = \{\bar{x} \mid \bar{x}_1 \ge 0, \bar{x}_2 \ge 0\}.$$

Consider the points:

$$\bar{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \in \Omega, \quad \bar{y} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \in \Omega.$$

Then, for $\alpha \in (0, 1)$,

$$\alpha\bar{x} + (1 - \alpha)\bar{y} = \begin{bmatrix} 2 - \alpha \\ 1 + \alpha \end{bmatrix}.$$

Evaluating f at this point:

$$f(\alpha\bar{x} + (1 - \alpha)\bar{y}) = (2 - \alpha)(1 + \alpha) = 2 + \alpha - \alpha^2.$$

On the other hand,

$$\alpha f(\bar{x}) + (1 - \alpha)f(\bar{y}) = 2.$$

For example, let $\alpha = \frac{1}{2}$, then:

$$f\left(\frac{1}{2}\bar{x} + \frac{1}{2}\bar{y}\right) = \frac{9}{4} > \frac{1}{2}f(\bar{x}) + \frac{1}{2}f(\bar{y}).$$

This shows that $f(\bar{x})$ is not convex over Ω. The answer is **no**.

References

[1] An Introduction To Optimization 4th Edition by Edwin K. P. Chong & Stanislaw H. Zak

[2] Convex Optimization by Stephen Boyd and Lieven Vandenberghe

[3] Introduction to Linear Optimization by Dimitris Bertsimas, John N. Tsitsiklis

[4] Wikipedia Integer Programming