國立交通大學
National Chiao Tung University

Department of Industrial Engineering and Management

Linear Programming

(5531)

Lecture 08 Convex Analysis

Prof. Sheng-I Chen

http://www.iem-omglab.nctu.edu.tw/OMG_LAB/

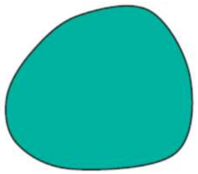
November 12, 2020

Agenda

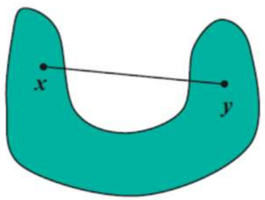
- Foundamentals of Convex Analysis
- Carathéodory’s Theorem
- Farkas’ Lemma
- Separation Theorem
- Strict Complementarity

Convex Sets

- **Definition:** A subset $S \in \mathbb{R}^n$ is called **convex** if for every x and y in S, S also contains *all* points on the line segment connecting x and y. That is $tx + (1 - t)y \in S$ for every $0 < t < 1$
- **Example:** The line segment connecting to any pair of points in the set is also contained in the set



- **Counter Example:** Existed a pair of points, such as x and y, their connecting line segment is not entirely in the set



Convex Sets

- **Examples**: All affine sets (including **0** and $\mathbb{R}^n$ itself) are convex

- **Examples**: Half-spaces are convex
  - The closed half-space $\{ x \mid \langle \boldsymbol{a},\boldsymbol{x} \rangle \leq b \}$ is convex
  - The open half-space $\{ x \mid \langle \boldsymbol{a},\boldsymbol{x} \rangle < b \}$ is convex

- **Theorem**: The intersection of an arbitrary collection of convex sets is convex

Convex Combination and Convex Set

- **Definition:** Given a finite set of **point** $z_1, z_2, . . . , z_n \in \mathbb{R}^m$, a point $z \in \mathbb{R}^m$ is called a **convex combination** of these points if:

$$z = \sum_{j=1}^{n} t_j z_j \quad , where\, t_j \geq 0 \quad and \quad \sum_{j=1}^{n} t_j = 1$$

  - The set of all convex combinations of two points is simply the line segment connecting them
  - It is called **strict convex combination** if *there is no t_j vanished*

- **Theorem:** A set C is **convex** *if only if* it contains *all convex combinations of points in* C

(Recap) Linear Combination & Affine Combination

- **Definition:** A vector $\mathbf{u}$ is said to be a **linear combination** of $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ in $\mathbb{R}^n$ if:

$$\mathbf{u} = \sum_{i=1}^{k} \lambda_i \mathbf{v_i} \qquad , \lambda_i \in R$$

- **Definition:** A vector $\mathbf{u}$ is said to be a **affine combination** of $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ in $\mathbb{R}^n$ if:

$$\mathbf{u} = \sum_{i=1}^{k} \lambda_i \mathbf{v_i} \quad \text{and} \quad \sum_{i=1}^{k} \lambda_i = 1 \qquad , \lambda_i \in R$$

Convex Functions and Convex Sets

- Convex functions are important sources of convex sets. One may determine a function is convex or not by checking its **level set** or **epigraph** $\{(x, f(x)) \mid f(x) \leq \alpha\}$

- **Definition**: Any continuous real-valued function f on $\mathbb{R}^n$ of open level sets $\{(x, f(x)) \mid f(x) \leq \alpha\}$ are convex if f is convex

- **Definition**: Let C be a convex subset of $\mathbb{R}^n$. We say that an extended function f: $C \rightarrow [-\infty, +\infty]$ is convex if the $epi(f)$ is a convex subset of $\mathbb{R}^{n+1}$

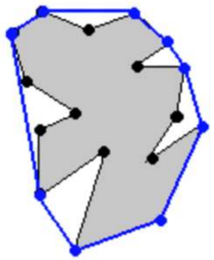
Another Definition for Convex Function

- **Proposition**: Let C be a nonempty convex subset of $\mathbb{R}^n$ and f: $\mathbb{R}^n \rightarrow \mathbb{R}$ be differentiable over an open set contained C. The function f is convex over C if and only if

$$f(z) \geq f(x) + \nabla f(x)^{\mathsf{T}} (z - x), \quad \forall\, x,z \in C$$

Convex Hull

- **Definition:** The set $conv(S)$ is called the **convex hull of S**:
  - $conv(S)$ *is* the *smallest convex set* containing each set S in $\mathbb{R}^n$ (note: not necessarily *S* is a convex set)
  - $conv(S)$ is the *intersection* of all convex sets containing S



- **Theorem:** The convex hull $conv(S)$ of a set S in $\mathbb{R}^n$ consists precisely of the set of all convex combinations of finite collections of points from S

Polytope and m-dimensional Simplex

- **Definition**: A convex hull of *finitely many* points is called a **polytope**
- **Definition**: Let $\{b_0, \ldots, b_m\}$ is affinely independent, its convex hull is called an **m-dimensional simplex** and $b_0, \ldots, b_m$ are the vertices of the simplex

  e.g. When m = 0,1,2 or 3 the simplex is a *point, line segment, triangle or tetrahedron*

  e.g. The point $\lambda_0 b_0 + \ldots + \lambda_m b_m$ with $\lambda_0 = \ldots = \lambda_m = 1 / (1+m)$ is called the *midpoint* or *barycenter* of the simplex

(Note) If $\sum_{i=1}^{k} \lambda_i \mathbf{v_i} = 0 \; and \sum_{i=1}^{k} \lambda_i = 0 \quad is \quad \lambda_i = 0 \quad for \quad all \quad i = 1,\ldots,k$, then the set of vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ is *affinely independent*

Another Definition of Polytope

- **Definition:** A **polyhedral set** (or a **polyhedron**) $P \subseteq \mathbb{R}^n$ is a set of points that satisfy a finite number of linear inequalities represented by $P = \{x \in \mathbb{R}^n : \mathbf{A}x \leq \mathbf{b}\}$

  ,where $\mathbf{A}$ is an m x n matrix and $\mathbf{b}$ is an n-vector

- **Definition:** A polyhedron is **bounded** if there exists an ω such that $P \subseteq \{x \in \mathbb{R}^n : -\omega \leq x \leq \omega, \forall j = 1,2,..,n\}$
  - *bounded polyhedral set* is called a **polytope**

A Polytope has finite vertices

Topological Properties

- **Definition**: A **norm** $\|.\|$ on $\mathbb{R}^n$ is a function that assigns a scalar $\|x\|$ to every $x \in \mathbb{R}^n$ and that has the following properties:
  1. $\|x\| \geq 0$ for all $x \in R^n$
  2. $\|\alpha\, x\| = |\alpha|.\ \|x\|$ for every scalar α and every $x \in R^n$
  3. $\|x\| = 0$ if an d only if $x = 0$
  4. $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in R^n$ (*Triangle Inequality*)

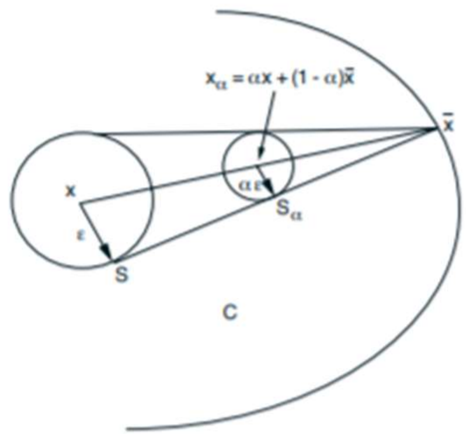
  Note: if x and y are *orthogonal*, then $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ (*Pythagorean Theorem*)

Closure, Interior and Relative Interior

- **Definition**: For any set C in $\mathbb{R}^n$. The **closure** of C, denoted as $cl(C)$, is expressed by:
  $cl(C) = \cap \{C + \varepsilon B \mid \varepsilon > 0\}$, where B is the *Euclidean unit ball* in R^n, $B = \{x \mid |x| \leq 1\}$
- **Proposition**: If the set C is convex, then the closure of C is also convex
- **Definition**: The **interior** of C, denoted as $int(C)$, is expressed by:
  $int(C) = \{x \mid \exists \varepsilon > 0,\ x + \varepsilon B \subset C\}$
- **Definition**: We say that x is a **relative interior point** of C if $x \in C$ and there exists an sphere centered at x $(x + \varepsilon B)$ such that $(x + \varepsilon B) \cap aff(C) \subset C$ (i.e., x is an interior point of C relative to the affine hull of C)
- **Definition**: The set of all relative interior points of C is called the **relative interior** of C, denoted as $ri(C)$, is formulated as:
  $ri(C) = \{x \in aff(C) \mid \exists \varepsilon > 0,\ (x + \varepsilon B) \cap aff(C) \subset C\}$

Line Segment Principle

- **Proposition**: Let C be a nonempty convex set. If $x \in ri(C)$ and $\bar{x} \in cl(C)$, then all points on the line segment connecting x and $\bar{x}$, except possible $\bar{x}$, belong to $ri(C)$



Bertsekas, Convex optimization theory

Ray / Direction / Cone / Convex Cone

- **Definition:** A **ray** is a *collection of points* of $\{x_0 + \lambda\, \mathrm{d} : \lambda \geq 0\}$, where x_0 is the *vertex of the ray*, and $\mathbf{d}$ is a nonzero vector called the **direction** of the ray
- **Definition:** A set C is said to be a **cone** if for all $x \in \mathrm{C}$ and $\lambda > 0$ we have $\lambda x \in \mathrm{C}$ (*Note that not all cones are convex*)
- **Definition:** A **Convex cone** C is a convex set with special property that $\lambda x \in \mathrm{C}$, $\forall\, x \in \mathrm{C}$ and $\lambda \geq 0$

Direction of Recession and Recession Cone

- **Definition**: Let C be a nonempty convex subset of $\mathbb{R}^n$. We say that a vector d is a **direction of recession** of C if $x + \alpha d \in C$ for all $x \in C$ and $\alpha \geq 0$
  - That is starting at any point x in C and going indefinitely along d and never cross the relatively boundary of C to points outside C
- **Definition**: The set of all directions of recession is a *cone containing the origin*, called the **recession cone** of C

Minkowski-Weyl Representation

- **Theorem**: A set P is polyhedral if and only if there is a nonempty finite set $\{v_1, \ldots, v_m\}$ and a finitely generated cone C such that $P = \text{conv}(\{v_1, \ldots, v_m\}) + C$, that is,

$$P = \{x \mid x = \sum_{j=1}^{m} m_j v_j + y,\ \sum_{j=1}^{m} m_j = 1, m_j \geq 0, j = 1, \ldots, m, y \in C\}$$

Note1: Thus, a polyhedral set can be generated via the convex combination of extreme points and cone
Note2: Finitely generated cone is convex
Note3: The vector sum of two convex set is convex
Note4: The term “finitely” means that the number of **half-spaces** to generate the cone is finite

Carathéodory's Theorem

- In the previous section, we showed that the convex hull of a set S can be constructed by forming *all* convex combinations of finite sets of points from S
- In 1907, Carathéodory showed that it is *not necessary to use all* finite sets. Instead, $m+1$ points are sufficient:
- **Theorem:** The convex hull $conv(S)$ of a set S in $\mathbb{R}^m$ consists of **all** convex combinations of **m + 1 (or less)** points from S:

$$conv(S) = \left\{ z = \sum_{j=1}^{m+1} t_j z_j : z_j \in S \quad , t_j \geq 0 \; for \quad all \quad j, and \quad \sum_j t_j = 1 \right\}$$

Proof Summary

H denotes the set of all convex combinations of $m+1$ points from S, and set S in $\mathbb{R}^m$. We want to show that $conv(S)$ is a subset of H

Let z be a point in $conv(S)$. By the previous theorem, $z = \sum_{j=1}^{n} t_j z_j$ (Note: n can be greater than $m+1$)

Let

$$A = [z_1 \quad z_2 \quad \dots \quad z_n] \quad and \quad x^* = \begin{bmatrix} t_1 \\ t_2 \\ . \\ . \\ t_n \end{bmatrix}$$

Then x^* is feasible for the following LP (Why? Number of constraints? Number of variables?):

$$\begin{aligned} &\text{maximize } c^T x \\ &\text{subject to } \; Ax = z \\ &\qquad\qquad\quad e^T x = 1 \\ &\qquad\qquad\quad x \geq 0 \end{aligned}$$

Since a feasible solution exists, then a basic feasible solution exists (Fundamental theorem of LP)

The basic variables are nonzero, and the number of basic variables is equal to the number of equality constraints.

In this LP, $m+1$ variables are nonzero.

The basic feasible solution corresponds to a convex combination of $m+1$ of the original n points ■

Farkas' Lemma

- **Lemma 10.5:** The system $Ax \leq b$ has no solutions *if and only if* there exists a vector y such that

$$A^T y = 0$$

$$y \geq 0$$

$$b^T y < 0$$

Farkas' Lemma

(1) $Ax \leq b$

(2) $A^T y = 0$

$y \geq 0$

$b^T y < 0$

If both (1) and (2) are true, then $0 = 0^T x = (A^T y)^T x = y^T Ax \leq y^T b = b^T y < 0$

"0 < 0" is a contradiction

Thus, either (1) or (2) can be true ∎

(Part of the Proof) (1) = False, (2) = True

Consider the LP

(P) $\quad maximize \ 0$

$\quad s.t \quad Ax \leq b$

It's dual is :

(D) $\quad minimize \quad b^T y$

$\quad s.t. \quad A^T y = 0$

$\quad y \geq 0$

If primal is infeasible then the dual is either infeasible or unbounded (*Fundamental theorem of LP*)

Clearly, the dual is always feasible (as there is a feasible solution with $y = 0$ and objective value $b^T y = 0$)

Therefore, if the primal is infeasible then the dual is unbounded

If the dual is unbounded, then there must existed a step direction Δy such that:

$A^T \Delta y = 0$ (i)

$\Delta y \geq 0$ (ii)

, and the new solution should decrease the objective value from zero :

$b^T \Delta y < 0$ (iii)

The solution Δy satisfies the (2) condition in the Farkas' lemma, when (1) has no solution ■

Variants of Farkas' Lemma

(1) has a solution:

$Ax \leq b$

(2) has a solution:

$A^T y = 0$

$y \geq 0$

$b^T y < 0$

Either (1) or (2) can be true

(1) has a solution:

$Ax \leq b$

$x \geq 0$

(2) has a solution:

$A^T y \geq 0$

$y \geq 0$

$b^T y < 0$

Either (1) or (2) can be true

(1) has a solution:

$Ax = b,\ x \geq 0$

(2) has a solution:

$A^T y \geq 0$

$b^T y < 0$

Either (1) or (2) can be true

The Separation Theorem

- **Definition:** A **halfspace** H of $\mathbb{R}^n$ to be any set given by a single linear inequality: $$H = \left\{x \in \mathbb{R}^n : \sum_{j=1}^{n} a_j x_j \leq b, (a_1, .., a_n) \neq 0\right\}$$
  - Note: If we allow $(a_1, a_2, \ldots, a_n) = 0$ then we call the set is a **generalized halfspace**
- **Definition:** A **polyhedron** is defined as the *intersection* of a finite collection of generalized halfspace. That is, a polyhedron is any set of the form: $$P = \left\{x \in \mathbb{R}^n : \sum_{j=1}^{n} a_{ij} x_j \leq b_i, i = 1, .., m\right\}$$
- **Theorem:** Let P *and* $\tilde{P}$ be two *disjoint* nonempty polyhedron in $\mathbb{R}^m$. There exist disjoint halfspaces H and $\tilde{H}$ such that $P \subset H$ *and* $\tilde{P} \subset \tilde{H}$

Proof of Separation Theorem

$Let \quad P, \tilde{P} \quad are \quad nonempty \quad polyhedron:$

$$P = \{x : Ax \le b\} \quad \tilde{P} = \left\{x : \tilde{A} x \le \tilde{b}\right\}$$

$$P \cap \tilde{P} = \varnothing$$

$$Thus, \begin{bmatrix} A \\ \tilde{A} \end{bmatrix} x \le \begin{bmatrix} b \\ \tilde{b} \end{bmatrix} has \quad no \quad solution$$

$By \quad Farkas' \quad lemma,$

$$A^T y + \tilde{A}^T \tilde{y} = 0$$

$$y \ge 0, \tilde{y} \ge 0$$

$$b^T y + \tilde{b}^T \tilde{y} < 0$$

$Let \quad H, \tilde{H} \quad are \quad halfspaces \quad defined \quad as:$

$$H = \{x:\ (A^T y)^T x \le b^T y\} \text{ and } \tilde{H} = \{x:\ (A^T y)^T x \ge -\tilde{b}^T \tilde{y}\}$$

$(we \quad need \quad to \quad show: P \subset H, \tilde{P} \subset \tilde{H}, and\ H \cap \tilde{H} = \varnothing)$

$(1) If\ x \in P, then \quad Ax \le b$

$$y^T Ax \le y^T b$$

$$(A^T y)^T x \le b^T y$$

$$x \in H$$

$$Thus, P \subset H$$

$(2)\ By \quad similar \quad analysis, we \quad can \quad show \quad \tilde{P} \subset \tilde{H}$

$(3)\ If \quad x \in H, (A^T y)^T x \le b^T y < -\tilde{b}^T \tilde{y} \quad (\because b^T y + \tilde{b}^T \tilde{y} < 0)$

$Thus, x \notin \tilde{H}$ ∎

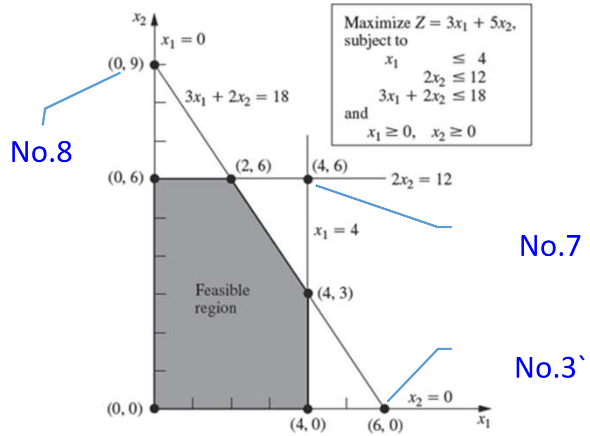
Strict Complementarity Theorems

- Recall in the Complementary Slackness:
  - (x^*, ω^*) *denotes an optimal solution to the primal and* (y^*, z^*) *denotes an optimal* solution to the dual. For each j *= 1, 2, . . . ,* n*, either* x_j^* = 0 *or* z_j^* = 0 *(or both are zeros) and, for each* i *= 1, 2, . . .,*m*,* either y_i^* = 0 or ω_i^* = 0 (or both are zeros)
- In this section, we will prove that there are optimal solutions for which exactly one member of each pair (x_j^*, z_j^*) vanishes and **exactly one** member from each pair (y_i^* , ω_i^*) vanishes. That is $x^* + z^* > 0$ and $y^* + \omega^* > 0$

(Recap) Complementary Slackness Theorem

- **Theorem 5.3.** $x = (x_1, x_2, \ldots, x_n)$ is a primal feasible solution and $y = (y_1, y_2, \ldots, y_m)$ is a dual feasible solution. x and y are optimal iff $x_j\,(z_j - c_j) = 0$ for $j = 1..n$ and $y_i\, x^s_i = 0$ for $i = 1..m$
- In No. 6, the dual solution is **feasible**, and thus the primal BFS is **optimal**
- In No. 3, both primal and dual are infeasible, but still satisfy the complementary slackness property
- Other BS solutions are either primal or dual feasible. If primal BFS is **not optimal**, then the corresponding dual solution is **not feasible**

	Primal Problem			Dual Problem	
No.	**Basic Solution**	**Feasible?**	**$Z = W$**	**Feasible?**	**Basic Solution**
1	(0, 0, 4, 12, 18)	Yes	0	No	(0, 0, 0, −3, −5)
2	(4, 0, 0, 12, 6)	Yes	12	No	(3, 0, 0, 0, −5)
3	(6, 0, −2, 12, 0)	No	18	No	(0, 0, 1, 0, −3)
4	(4, 3, 0, 6, 0)	Yes	27	No	$\left(-\frac{9}{2}, 0, \frac{5}{2}, 0, 0\right)$
5	(0, 6, 4, 0, 6)	Yes	30	No	$\left(0, \frac{5}{2}, 0, -3, 0\right)$
6	(2, 6, 2, 0, 0)	Yes	36	Yes	$\left(0, \frac{3}{2}, 1, 0, 0\right)$
7	(4, 6, 0, 0, −6)	No	42	Yes	$\left(3, \frac{5}{2}, 0, 0, 0\right)$
8	(0, 9, 4, −6, 0)	No	45	Yes	$\left(0, 0, \frac{5}{2}, \frac{9}{2}, 0\right)$



Strict Complementarity

$$\begin{array}{ll} \text{Primal} & \\ \text{maximize} & c^T x \\ \text{subject to} & Ax + \omega = 0 \\ & x, \omega \ge 0 \end{array} \qquad \begin{array}{ll} \text{Dual} & \\ \text{minimize} & b^T y \\ \text{subject to} & A^T y - z = c \\ & y, z \ge 0 \end{array}$$

- **Theorem 10.6.** If both the primal and the dual have feasible solutions, then there exists a primal feasible solution $(\bar{x}, \bar{\omega})$ and a dual feasible solution $(\bar{y}, \bar{z})$ such that $\bar{x} + \bar{z} > 0$ and $\bar{y} + \bar{\omega} > 0$

- **Theorem 10.7.** If both the primal and the dual have optimal solutions, then there exists a primal optimal solution (x^*, ω^*) and a dual optimal solution(y^*, z^*) such that (y^*, z^*) $x^* + z^* > 0$ and $y^* + \omega^* > 0$

Proof of theorem 10.6:

- (*Case 1*) If there is a feasible solution for which $\bar{x}_j > 0$, then $\bar{x}_j + \bar{z}_j > 0$
- (*Case 2*) Let's consider another situation when $x_j = 0$ for **every primal feasible solution**. Then, the following LP is also feasible since its constraints are the same as for the original primal. Additionally, the following LP has an optimal solution with objective value **0** (why?):

$$\begin{aligned} \max\ & \boldsymbol{x_j} \\ s.t.\ & Ax \le b \\ & x \ge 0 \end{aligned}$$

, and its corresponding dual is:

$$\begin{aligned} & \min b^T y \\ & s.t. \\ & A^T y \ge \boldsymbol{e_j} \\ & y \ge 0 \end{aligned}$$

Dual

$$\begin{aligned} &\min imize \quad b^T y \\ &subject\ to \quad A^T y - z = c \\ &\qquad\qquad\qquad y, z \geq 0 \end{aligned}$$

(Continue the proof)

- By strong duality, the dual (note: this dual is not the dual of the original primal) has an optimal solution (y', z'), and thus we have $A^T y' - z' = e_j$ and $y', z' \geq 0$ (1)
- Let (y, z) be any feasible solution to the dual of the original primal (at the right top corner)
- The properties of y' and z' in (1) implies that $y + y'$ is feasible for dual and its slack is $z + z' + e_j$
- Clearly, this dual feasible solution have the jth component of slack variables is at least 1
- Combining the analysis of cases 1 and 2, for each *j*, there exists a primal feasible ($x^{(j)}$, $\omega^{(j)}$) and dual feasible solution ($y^{(j)}$, $z^{(j)}$) such that $x_j^{(j)} + z_j^{(j)} > 0$
- To complete the proof, we take a strict convex combination of these solutions to get one solution that works for all j
- Since the feasible region is convex, the convex combinations preserve primal and dual feasibility
- Since the convex combination is strict, it follows that every primal variable and its dual slack add to a strictly positive number as every dual variable and its primal slack ■

Homework

- (Required. Handwriting) Exercise 10.4 Find a strictly complementary solution to the following linear programming problem and its dual.
- (Self-practice. Computational exercise) Find the vertexes of convex hull of the integer program solution
  - Description: Given a set of IP solutions, $\mathbf{S}=\{x \in B^5: 79x_1 + 53x_2 + 53x_3 + 45x_4 + 45x_5 \leq 178\}$, (1) find S; (2) find the vertexes of $conv(\mathbf{S})$
  - (Hint) The procedure to find the vertexes of $conv(\mathbf{S})$ is as follows:

1. Let **S** be the set of IP feasible solutions;
2. Let **v** be the set of vertexes discovered from the last step;
3. **For** *k* in **S**
4. **If** x_k <> $\lambda x_i + (1-\lambda) x_j$, for any *i,j* in **S then**
5. x_k is a vertex point;
6. v ← v∪x_k;
7. **End if**
8. **End for**

(Note: the complexity of this step is **O**(n /**S**/^{3}), where n is the number of decision variables for the IP and /**S**/ is the number of IP feasible solutions)