

Notations:

LI: Linearly independent

LD: Linearly dependent

 $\mathbf{x}, \mathbf{d}, \mathbf{b}$, etc, that is characters in boldface represent (column) vectors $S = \text{Fea}(LPP) = \{\mathbf{x} \in \mathbb{R}^n : A_{m \times n} \mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$.

Observation 5: Suppose if a LPP has an unbounded feasible region, then there exists a vector $\mathbf{d} \neq \mathbf{0}$ such that starting from any point of the feasible region if you move in the positive direction of \mathbf{d} , then you will always remain in the feasible region.

That is for any $\mathbf{x} \in \text{Fea}(LPP)$, $\mathbf{x} + \alpha \mathbf{d} \in \text{Fea}(LPP)$ for all $\alpha \geq 0$.Then $\mathbf{d} \neq \mathbf{0}$ is called a **direction** of $S = \text{Fea}(LPP)$.Throughout our discussion, \mathbf{d} will denote a column vector given by $\mathbf{d} = [d_1, \dots, d_n]^T$.

Definition: Given a nonempty convex set S , $S \subset \mathbb{R}^n$, $\mathbf{d} \neq \mathbf{0}$ is called a **direction** of S if for all $\mathbf{x} \in S$, $\mathbf{x} + \alpha \mathbf{d} \in S$ for all $\alpha \geq 0$.

From the definition it is clear that if \mathbf{d} is a direction of a convex set S , then for all $\gamma > 0$, $\mathbf{x} + \alpha \mathbf{d} = \mathbf{x} + (\frac{\alpha}{\gamma}) \gamma \mathbf{d} \in S$ for all $\alpha > 0$ $\Rightarrow \gamma \mathbf{d}$ is a direction for all $\gamma > 0$.

Two directions $\mathbf{d}_1, \mathbf{d}_2$ of S are said to be distinct if $\mathbf{d}_1 \neq \gamma \mathbf{d}_2$ for any $\gamma > 0$ (or equivalently $\mathbf{d}_2 \neq \beta \mathbf{d}_1$ for any $\beta > 0$).

Example 2: (revisited) Consider the problem,Min $-x + 2y$

subject to

 $x + 2y \geq 1$ $-x + y \leq 1$, $x \geq 0, y \geq 0$.Note that $\mathbf{d}_1 = [1, 1]^T$, $\mathbf{d}_1 = [2, 2]^T, \dots$ are all equal as directions of $\text{Fea}(LPP)$.Similarly $\mathbf{d}_1 = [1, 0]^T$, $\mathbf{d}_1 = [2, 0]^T, \dots$ are all equal as directions of $\text{Fea}(LPP)$.Whereas $\mathbf{d}_1 = [1, 1]^T$, $\mathbf{d}_2 = [1, 0]^T$ give two distinct directions.**Result:** The set of all directions of $S = \{\mathbf{x} \in \mathbb{R}^n : A_{m \times n} \mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ is given by $D = \{\mathbf{d} \in \mathbb{R}^n : \mathbf{d} \neq \mathbf{0}, \quad A_{m \times n} \mathbf{d} \leq \mathbf{0}, \quad \mathbf{d} \geq \mathbf{0}\}$ or $D = \{\mathbf{d} \in \mathbb{R}^n : \mathbf{d} \neq \mathbf{0}, \quad \mathbf{a}_i^T \mathbf{d} \leq 0, \text{ for all } i = 1, 2, \dots, m, \quad \mathbf{d} \geq \mathbf{0}\}$.**Proof:** If $\mathbf{d} \in D$ and $\mathbf{x} \in S$, then(1) $\mathbf{x} + \alpha \mathbf{d} \geq \mathbf{0}$ for all $\alpha \geq 0$ since $\mathbf{x} \geq \mathbf{0}$ and $\mathbf{d} \geq \mathbf{0}$.(2) Also $A(\mathbf{x} + \alpha \mathbf{d}) = A\mathbf{x} + \alpha A\mathbf{d} \leq \mathbf{b}$, for all $\alpha \geq 0$ since $A\mathbf{x} \leq \mathbf{b}$, $A\mathbf{d} \leq \mathbf{0}$.From (1) and (2), $\mathbf{x} + \alpha \mathbf{d} \in S$ for all $\alpha \geq 0$.Hence if $\mathbf{d} \in D$ then \mathbf{d} is a direction of S . (*)For the converse, let \mathbf{d} be such that it does not belong to D .Then either $d_i < 0$ for some $i = 1, 2, \dots, n$, (1)or $(A\mathbf{d})_j = \mathbf{a}_j^T \mathbf{d} > 0$ for some $j = 1, 2, \dots, m$. (2)(1) If $d_i < 0$ for some $i = 1, 2, \dots, n$

then given any $\mathbf{x} \in S$ there exists $\alpha > 0$ sufficiently large such that, $x_i + \alpha d_i < 0$,

$\Rightarrow (\mathbf{x} + \alpha \mathbf{d})$ does not belong to S for all such α

$\Rightarrow \mathbf{d}$ is not a direction of S .

(2) If $(A\mathbf{d})_j = \mathbf{a}_j^T \mathbf{d} > 0$ for some $j = 1, 2, \dots, m$,

then given any $\mathbf{x} \in S$ there exists $\alpha > 0$ sufficiently large such that

$(A\mathbf{x})_j + \alpha(A\mathbf{d})_j > b_j$, hence $(\mathbf{x} + \alpha \mathbf{d})$ does not belong to S for all such α ,

$\Rightarrow \mathbf{d}$ is not a direction of S .

Hence if \mathbf{d} does not belong to D then \mathbf{d} cannot be a direction of S .

Or (contrapositive) if \mathbf{d} is a direction of S then $\mathbf{d} \in D$. (**)

(*) and (**) together gives the required result.

Remark: Note that the set of all directions of $S = \text{Fea}(LPP)$ is a convex set.

In fact if $\mathbf{d}_1, \mathbf{d}_2$ are two directions of S , then $\mathbf{d} = \alpha \mathbf{d}_1 + \beta \mathbf{d}_2$ will again be a direction of S , for any α, β nonnegative (as long as both α, β are not equal to zero or $\alpha + \beta \neq 0$).

Definition: A direction \mathbf{d} of S is called an **extreme direction** of S , if it cannot be written as a positive linear combination of two distinct directions of S , that is, if

$\mathbf{d} = \alpha \mathbf{d}_1 + \beta \mathbf{d}_2$, for $\alpha, \beta > 0$ and $\mathbf{d}_1, \mathbf{d}_2 \in D$ then $\mathbf{d}_1 = \gamma \mathbf{d}_2$ for some $\gamma > 0$.

It is clear that if D denotes the set of all directions of S

(which will be the empty set if S is bounded) then $D' = \{\mathbf{d} \in \mathbb{R}^n : \mathbf{d} \geq \mathbf{0}, A\mathbf{d} \leq \mathbf{0}, \sum_i d_i = 1\}$ is a set of all distinct directions of S .

Also each $\mathbf{d} \in D$ is of the form $\mathbf{d} = \alpha \mathbf{d}'$ for some $\mathbf{d}' \in D'$ and $\alpha = \sum_i d_i (> 0)$.

Note that D' can be written as

$$D' = \left\{ \mathbf{d} \in \mathbb{R}^n : \mathbf{d} \geq \mathbf{0}, \begin{bmatrix} A & & \\ 1 & 1, \dots, 1 \\ -1 & -1, \dots, -1 \end{bmatrix} \mathbf{d} \leq \begin{bmatrix} \mathbf{0} \\ 1 \\ -1 \end{bmatrix} \right\}.$$

The set D' now looks exactly like the feasible region of an LPP, hence if D' is nonempty then D' has at least one extreme point (why?).

Result: \mathbf{d} is an extreme direction of S if and only if $\mathbf{d}' = \frac{\mathbf{d}}{\sum_i d_i}$ is an extreme point of D' .

Proof: Let $\mathbf{d}, \mathbf{d}_1, \mathbf{d}_2 \in D$ be such that $\mathbf{d} = (d_1, \dots, d_n)^T, \mathbf{d}_1 = (d_{11}, \dots, d_{1n})^T$ and $\mathbf{d}_2 = (d_{21}, \dots, d_{2n})^T$.

Since \mathbf{d}, \mathbf{d}_1 and \mathbf{d}_2 are all nonnegative and nonzero vectors, $\sum_i d_i, \sum_i d_{1i}, \sum_i d_{2i} > 0$.

Let $\mathbf{d}' = \frac{\mathbf{d}}{\sum_i d_i}, \mathbf{d}'_1 = \frac{\mathbf{d}_1}{\sum_i d_{1i}}, \mathbf{d}'_2 = \frac{\mathbf{d}_2}{\sum_i d_{2i}}$, then $\mathbf{d}', \mathbf{d}'_1, \mathbf{d}'_2 \in D'$.

Let $\alpha, \beta > 0$ be such that

$$\mathbf{d} = \alpha \mathbf{d}_1 + \beta \mathbf{d}_2, \quad (*)$$

$$\iff \frac{\mathbf{d}}{\sum_i d_i} = \alpha \left(\frac{\sum_i d_{1i}}{\sum_i d_i} \right) \frac{\mathbf{d}_1}{\sum_i d_{1i}} + \beta \left(\frac{\sum_i d_{2i}}{\sum_i d_i} \right) \frac{\mathbf{d}_2}{\sum_i d_{2i}}, \quad (**)$$

$$\iff \mathbf{d}' = \lambda \mathbf{d}'_1 + (1 - \lambda) \mathbf{d}'_2, \quad (***)$$

where $\lambda = \alpha \left(\frac{\sum_i d_{1i}}{\sum_i d_i} \right), 0 < \lambda < 1$.

(Since $\mathbf{d} = \alpha \mathbf{d}_1 + \beta \mathbf{d}_2, \sum_i d_i = \alpha(\sum_i d_{1i}) + \beta(\sum_i d_{2i})$).

Since $\mathbf{d}_1, \mathbf{d}_2 \in D$ are distinct as directions $\iff \mathbf{d}'_1 \neq \mathbf{d}'_2$,

from (*) and (***) it follows that

\mathbf{d} is not (or is) an extreme direction of $S \iff \mathbf{d}'$ is not (or is) an extreme point of D' .

Remark: Hence the number of distinct extreme directions of S is finite (why?).

Also since $D' \subset \mathbb{R}^n$ is a polyhedral set with non n negativity constraints (like $Fea(LPP) = S$), if $D' \neq \phi$, then D' must have atleast one extreme point (not proved as yet).

Hence if $Fea(LPP) = S$ is unbounded S must have atleast one extreme direction.

Also the extreme directions of S which are also extreme points of D' (after suitable normalization) will lie on n LI hyperplanes defining D' .

Since any $\mathbf{d} \in \mathbb{R}^n$, $\mathbf{d} \neq \mathbf{0}$ cannot be orthogonal to n LI vectors, so \mathbf{d} cannot lie on n LI hyperplanes of the $(m+n)$ hyperplanes given by,

$\{\mathbf{d} \in \mathbb{R}^n : \mathbf{a}_i^T \mathbf{d} = 0\}$ for $i = 1, 2, \dots, m$, and $\{\mathbf{d} \in \mathbb{R}^n : -\mathbf{e}_j^T \mathbf{d} = 0\}$ for $j = 1, 2, \dots, n$.

So if $\mathbf{d} \in D'$, is an extreme direction of S (or an extreme point of D'), then it should lie on $(n-1)$ LI hyperplanes of the above mentioned $(m+n)$ hyperplanes, and the hyperplane $\{\mathbf{d} \in \mathbb{R}^n : [1, 1, \dots, 1]\mathbf{d} = 1\}$ on which \mathbf{d} must necessarily lie (since $\mathbf{d} \in D'$). This gives a collection of n LI hyperplanes, on which \mathbf{d} lies.

So any $\mathbf{d} \in D$, which lies on $(n-1)$ LI hyperplanes out of the $(m+n)$ hyperplanes given by

$\{\mathbf{d} \in \mathbb{R}^n : \mathbf{a}_i^T \mathbf{d} = 0\}$ for $i = 1, 2, \dots, m$, and $\{\mathbf{d} \in \mathbb{R}^n : -\mathbf{e}_j^T \mathbf{d} = 0\}$ for $j = 1, 2, \dots, n$,

is an extreme direction of S .

Exercise: Check that if a $\mathbf{d} \in D$ lies on $(n-1)$ LI hyperplanes (out of the $(m+n)$ defining hyperplanes of D) given by $\{H_1, \dots, H_{n-1}\}$, then $\{H, H_1, \dots, H_{n-1}\}$ is LI where $H = \{\mathbf{d} \in \mathbb{R}^n : [1, 1, \dots, 1]\mathbf{d} = 1\}$.

Example 2: (revisited) Consider the problem,

Min $-x + 2y$

subject to

$x + 2y \geq 1$

$-x + y \leq 1$,

$x \geq 0, y \geq 0$.

Note that here the set of all directions of S is given by

$D = \{\mathbf{d} \in \mathbb{R}^2 : [-1, -2]\mathbf{d} \leq 0, [-1, 1]\mathbf{d} \leq 0, \mathbf{d} \geq \mathbf{0}\}$.

Also if $\mathbf{d} \in D$ is an extreme direction of S then it has to lie on exactly one of the hyperplanes given by

(i) $\{\mathbf{d} \in \mathbb{R}^2 : [-1, -2]\mathbf{d} = 0\}$,

(ii) $\{\mathbf{d} \in \mathbb{R}^2 : [-1, 1]\mathbf{d} = 0\}$,

(iii) $\{\mathbf{d} \in \mathbb{R}^2 : d_1 = 0\}$,

(iv) $\{\mathbf{d} \in \mathbb{R}^2 : d_2 = 0\}$.

Note that there exists no $\mathbf{d} \geq \mathbf{0}$, $\mathbf{d} \neq \mathbf{0}$ such that $[-1, -2]\mathbf{d} = 0$.

Also if $\mathbf{d} \geq \mathbf{0}$, satisfies the condition $d_1 = 0$, then $[-1, 1]\mathbf{d} \leq 0$ cannot be satisfied, hence such a \mathbf{d} does not belong to D .

Hence $\mathbf{d} \in D$, is an extreme direction of S if it lies on either the hyperplane

$\{\mathbf{d} \in \mathbb{R}^2 : [-1, 1]\mathbf{d} = 0\}$, or in $\{\mathbf{d} \in \mathbb{R}^2 : d_2 = 0\}$.

Hence $\mathbf{d}' = [1, 1]^T$ and any positive scalar multiple of \mathbf{d}' (they are all same as directions), and $\mathbf{d}'' = [1, 0]^T$ and any positive scalar multiple of \mathbf{d}'' , are the only possible extreme directions of $S = Fea(LPP)$ of the LPP given above.

Theorem: If $S = Fea(LPP) = \{\mathbf{x} \in \mathbb{R}^n : A_{m \times n} \mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ is nonempty, then S has atleast one extreme point.

Proof: Consider $\mathbf{x}_0 \in S$. If \mathbf{x}_0 is an extreme point of S , then done.

If not, then let \mathbf{x}_0 lie on **exactly**, $0 \leq k < n$, LI defining hyperplanes of S . Also there exists $\mathbf{y}_1, \mathbf{y}_2$ **distinct** elements of S such that \mathbf{x}_0 lies strictly in between and on the line segment joining $\mathbf{y}_1, \mathbf{y}_2$,

that is,

there exists $0 < \lambda < 1$, such that $\mathbf{x}_0 = \lambda \mathbf{y}_1 + (1 - \lambda) \mathbf{y}_2$, $\mathbf{y}_1, \mathbf{y}_2 \in S$, $\mathbf{y}_1 \neq \mathbf{y}_2$.

Let the k LI defining hyperplanes on which \mathbf{x}_0 lies WLOG be H_1, \dots, H_k , and let the corresponding normals be $\tilde{\mathbf{a}}_j$, $j = 1, 2, \dots, k$.

Then $\{\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_k\}$ is LI and $\tilde{\mathbf{a}}_i^T \mathbf{x}_0 = \tilde{b}_i$, for $i = 1, 2, \dots, k$.

Also note that each of $\mathbf{y}_1, \mathbf{y}_2$ also lies on these k , LI hyperplanes (we have seen this earlier also while proving the equivalence of the definition of corner points and extreme points), that is,

$\tilde{\mathbf{a}}_i^T \mathbf{y}_1 = \tilde{b}_i$, and $\tilde{\mathbf{a}}_i^T \mathbf{y}_2 = \tilde{b}_i$, for all $i = 1, 2, \dots, k$.

If $\mathbf{d} = \mathbf{y}_2 - \mathbf{y}_1$, then $\mathbf{d} \neq \mathbf{0}$ and \mathbf{d} is orthogonal to the normals of each of the k hyperplanes on which \mathbf{x}_0 lies, that is for all $i = 1, \dots, k$,

$\tilde{\mathbf{a}}_i^T \mathbf{d} = \tilde{\mathbf{a}}_i^T (\mathbf{y}_2 - \mathbf{y}_1) = \tilde{b}_i - \tilde{b}_i = 0$.

Consider feasible points of the form $(\mathbf{x}_0 \pm \alpha \mathbf{d})$, that is points of S obtained by moving in the positive and negative direction of \mathbf{d} starting from \mathbf{x}_0 .

Then note that for all $i = 1, 2, \dots, k$, $\tilde{\mathbf{a}}_i^T (\mathbf{x}_0 \pm \alpha \mathbf{d}) = \tilde{\mathbf{a}}_i^T \mathbf{x}_0 = \tilde{b}_i$ for all $\alpha \in \mathbb{R}$ (*).

Since $\mathbf{x}_0 \geq \mathbf{0}$ and $\mathbf{d} \neq \mathbf{0}$, there exists $\alpha > 0$ large,

such that either $\mathbf{x}_0 + \alpha \mathbf{d}$ does not belong to S or $\mathbf{x}_0 - \alpha \mathbf{d}$ does not belong to S .

Let us assume WLOG that $\mathbf{x}_0 - \alpha \mathbf{d}$ does not belong to S for α large,

and let $\gamma = \max\{\alpha > 0 : \mathbf{x}_0 - \alpha \mathbf{d} \in S\}$, then $\gamma > 0$ (since $\mathbf{y}_1 \in S$ and $\mathbf{y}_1 = \mathbf{x}_0 - (1 - \lambda) \mathbf{d}$ where $1 - \lambda > 0$).

Also, $\mathbf{x}_1 = \mathbf{x}_0 - \gamma \mathbf{d} \in S$ lies in each of the k LI hyperplanes on which \mathbf{x}_0 lies (by (*)) and also lies in atleast one more hyperplane say H_0 , which obstructs further movement along the direction of $-\mathbf{d}$, starting from \mathbf{x}_0 .

Let the normal of H_0 be $\tilde{\mathbf{a}}_0$,

then $\tilde{\mathbf{a}}_0^T (\mathbf{x}_0 - \gamma \mathbf{d}) = \tilde{b}_0$, but $\tilde{\mathbf{a}}_0^T (\mathbf{x}_0 - \alpha \mathbf{d}) > \tilde{b}_0$ for all $\alpha > \gamma$. (**)

Check that the hyperplanes $H_0, H_1, H_2, \dots, H_k$, are LI.

If suppose not, then suppose the set $\{\tilde{\mathbf{a}}_0, \tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_k\}$ is LD.

Since $\{\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_k\}$ is LI it implies

$\tilde{\mathbf{a}}_0$ can be written as a linear combination of $\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_k$,

$\Rightarrow \mathbf{d}$ is also orthogonal to $\tilde{\mathbf{a}}_0$, that is $\tilde{\mathbf{a}}_0^T \mathbf{d} = \mathbf{0}$.

$\Rightarrow \tilde{\mathbf{a}}_0^T (\mathbf{x}_0 - \alpha \mathbf{d}) = \tilde{\mathbf{a}}_0^T \mathbf{x}_0 = \tilde{\mathbf{a}}_0^T (\mathbf{x}_0 - \gamma \mathbf{d}) = \tilde{b}_0$ for all $\alpha \in \mathbb{R}$, which contradicts (**).

Hence the hyperplanes $H_0, H_1, H_2, \dots, H_k$ are LI, and we obtain an $\mathbf{x}_1 \in S$, which lies on atleast $(k + 1)$, LI hyperplanes defining S . If \mathbf{x}_1 is an extreme point, then again done. If not then continue as before starting now from the point \mathbf{x}_1 . Hence after at most $(n - k)$ steps we will find a feasible point which lies on exactly n LI hyperplanes defining S , and hence is an extreme point of S .

Remark: Note that the above result is not necessarily true for all polyhedral sets.

For example take any single half space, or say a straight line in \mathbb{R}^n , which are polyhedral sets, but does not have any extreme point.

The theorem works for $Fea(LPP)$ because of the nonnegativity constraints, that is because $Fea(LPP)$ is given a supply of n LI hyperplanes, among the $(m+n)$ hyperplanes defining S .

Exercise: Can you find a nonempty polyhedral set S , $S \subset \mathbb{R}^3$ which has two defining hyperplanes but does not have any extreme point.

Exercise: Can you find a nonempty polyhedral set S , $S \subset \mathbb{R}^3$ which has three LI defining hyperplanes (not necessarily the nonnegativity constraints) but does not have any extreme point.

Definition: Given S , a nonempty subset of \mathbb{R}^n , and $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k \in S$, $\sum_{i=1}^k \lambda_i \mathbf{x}_i$, is called a convex combination of $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$,

where $0 \leq \lambda_i \leq 1$ for all $i = 1, 2, \dots, k$, and $\sum_{i=1}^k \lambda_i = 1$.

Result: Given $\phi \neq S \subset \mathbb{R}^n$, S is a convex set if and only if for all $k \in \mathbb{N}$, the convex combination of any k elements of S is again an element of S .

Proof: ‘If part’ is obvious, follows from the definition of convex sets.

To show the ‘Only if’ part, that is to show that if S is a convex set then the convex combination of any collection of finitely many elements of S belongs to S .

We will prove this by induction on k .

Since S is convex so the result is true for $k = 2$.

Assume that the convex combination of any $n \leq k$ points of S is in S , to show that the convex combination of any $(k + 1)$ points of S is in S .

Let $\mathbf{x} = \sum_{i=1}^{k+1} \lambda_i \mathbf{x}_i$, where $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{k+1} \in S$,

$0 \leq \lambda_i \leq 1$, for all $i = 1, 2, \dots, k + 1$ and $\sum_{i=1}^{k+1} \lambda_i = 1$.

then $1 - \lambda_{k+1} = \sum_{i=1}^k \lambda_i$ and $\mathbf{x} = (1 - \lambda_{k+1}) \left(\sum_{i=1}^k \frac{\lambda_i \mathbf{x}_i}{\sum_{i=1}^k \lambda_i} \right) + \lambda_{k+1} \mathbf{x}_{k+1}$.

Note that $\sum_{i=1}^k \frac{\lambda_i \mathbf{x}_i}{\sum_{i=1}^k \lambda_i} \in S$ by induction hypothesis and $\mathbf{x}_{k+1} \in S$.

Hence \mathbf{x} which is now expressed as a convex combination of two elements of S , belongs to S .

Theorem: (Representation Theorem) Let $S = \{\mathbf{x} \in \mathbb{R}^n : A_{m \times n} \mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\} \neq \emptyset$ such that $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ are the extreme points of S and $\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_r$ are the distinct extreme directions of S (the set of directions is empty if S is bounded) then $\mathbf{x} \in S$ if and only if

$\mathbf{x} = \sum_{i=1}^k \lambda_i \mathbf{x}_i + \sum_{j=1}^r \mu_j \mathbf{d}_j$ for some λ_i ’s and μ_j ’s

where $0 \leq \lambda_i \leq 1$ for all $i = 1, 2, \dots, k$, $\sum_i \lambda_i = 1$, and $\mu_j \geq 0$, for all $j = 1, 2, \dots, r$.

That is, $\mathbf{x} \in S$ if and only if it can be written as a convex combination of the extreme points of S plus a nonnegative linear combination of the extreme directions of S .

Proof: The ‘If part’ can be verified easily.

That is, if \mathbf{x}_0 is of the form

$$\mathbf{x}_0 = \sum_{i=1}^k \lambda_i \mathbf{x}_i + \sum_{j=1}^r \mu_j \mathbf{d}_j$$

where $0 \leq \lambda_i \leq 1$, for all $i = 1, 2, \dots, k$, $\sum_{i=1}^k \lambda_i = 1$, and $\mu_j \geq 0$ for all $j = 1, 2, \dots, r$, then to see that $\mathbf{x}_0 \in S$.

$\mathbf{x}_0 \geq \mathbf{0}$ is obvious, since each of the \mathbf{x}_i ’s and \mathbf{d}_j ’s are nonnegative vectors, and all that λ_i ’s and μ_j ’s are nonnegative.

Since $\mathbf{x}_i \in S$, $A\mathbf{x}_i \leq \mathbf{b}$ for all $i = 1, 2, \dots, k$. (1)

Since $\mathbf{d}_j \in D$, $A\mathbf{d}_j \leq \mathbf{0}$ for all $j = 1, 2, \dots, r$. (2)

Also

$$\begin{aligned} A\mathbf{x}_0 &= \sum_{i=1}^k \lambda_i A\mathbf{x}_i + \sum_{j=1}^r \mu_j A\mathbf{d}_j \\ &\leq \sum_{i=1}^k \lambda_i A\mathbf{x}_i \quad \text{follows from (2) and from } \mu_j \geq 0 \text{ for all } j \\ &\leq \sum_{i=1}^k \lambda_i \mathbf{b} \quad \text{follows from (1) and from } \lambda_i \geq 0 \text{ for all } i \\ &= \mathbf{b}, \quad \text{since } \sum_{i=1}^k \lambda_i = 1. \end{aligned}$$

Hence \mathbf{x}_0 satisfies the conditions $A\mathbf{x} \leq \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$, which implies $\mathbf{x}_0 \in S$.

‘Only if’ part.

Let us assume that S is unbounded and let \mathbf{x}_0 be an arbitrary element of S .

If \mathbf{x}_0 is an extreme point of S , WLOG let us assume $\mathbf{x}_0 = \mathbf{x}_1$,

then $\mathbf{x}_0 = 1.\mathbf{x}_1 + 0.\mathbf{x}_2 + \dots + 0.\mathbf{x}_k + 0.\mathbf{d}_1 + \dots + 0.\mathbf{d}_r$

which is a convex combination of the extreme points of S and nonnegative linear combination of extreme directions of S .

If not, that is if \mathbf{x}_0 is not an extreme point of S then choose an $M > 0$, large such that $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_k \in \bar{S}$ where $\bar{S} = \{\mathbf{x} \in S : \sum_{i=1}^n x_i \leq M\}$, and none of the extreme points of S or \mathbf{x}_0 lies on the newly added hyperplane $H_0 = \{\mathbf{x} \in \mathbb{R}^n : \sum_{i=1}^n x_i = M\}$.

Note that \bar{S} is bounded.

Since \bar{S} has $m + n + 1$ constraints of which $(m + n)$ come from S , so all the extreme points of S are also extreme points of \bar{S} (since they lie on n LI hyperplanes defining S), but some new extreme points may have been added to \bar{S} due to the addition of the new hyperplane H_0 .

Since \mathbf{x}_0 is not an extreme point of S or \bar{S} , let us assume that it lies on exactly k ($0 \leq k < n$) LI hyperplanes defining S . Also there exists a line segment $L_{\mathbf{x}_0}$ (with \mathbf{x}_0 sitting strictly in between) with boundary points $\mathbf{y}_1, \mathbf{y}_2$ totally contained in \bar{S} .

Note that both $\mathbf{y}_1, \mathbf{y}_2$ also lies on the k LI hyperplanes on which \mathbf{x}_0 lies. Let $\mathbf{d} = \mathbf{y}_1 - \mathbf{y}_2$, then $\mathbf{x}_0 + \alpha \mathbf{d} \in \bar{S}$ for $\alpha > 0$ sufficiently small.

Let $\gamma = \max\{\alpha : \mathbf{x}_0 + \alpha \mathbf{d} \in \bar{S}\}$ (there exists such a $\gamma > 0$ since \bar{S} is bounded).

Let $\mathbf{y} = \mathbf{x}_0 + \gamma \mathbf{d}$, then \mathbf{y} lies on atleast $(k + 1)$ LI hyperplanes defining \bar{S} of which k (are from defining hyperplanes of S) in common with $\mathbf{x}_0, \mathbf{y}_1, \mathbf{y}_2$.

Now by starting with \mathbf{y} and repeating the above process, after atmost $n - k - 1$ steps we will be able to find an extreme point of \bar{S} , call it \mathbf{x}_{i_1} such that that this extreme point lies on n lie hyperplanes defining \bar{S} of which k are common with \mathbf{x}_0 .

Consider the line segment joining \mathbf{x}_{i_1} and \mathbf{x}_0 (all points on this line segment will be in \bar{S} since it is a convex set) and extend it further from \mathbf{x}_0 in the positive direction of the vector $\mathbf{d}_0 = \mathbf{x}_0 - \mathbf{x}_{i_1}$ (you will be able to extend it further from \mathbf{x}_0 since otherwise if there is any obstruction of movement at \mathbf{x}_0 , then it must be by a hyperplane which is LI to the first k hyperplanes on which \mathbf{x}_0 lies, which will contradict that \mathbf{x}_0 lies on exactly k LI hyperplanes defining \bar{S}).

Let $\beta = \max\{\alpha : \mathbf{x}_{i_1} + \alpha \mathbf{d}_0 \in \bar{S}\}$ (there exists such a $\beta > 1$ since \bar{S} is bounded) and let $\mathbf{y}_0 = \mathbf{x}_{i_1} + \beta \mathbf{d}_0$, then \mathbf{y}_0 lies on atleast $(k + 1)$ LI hyperplanes defining \bar{S} of which k are from S , in common with $\mathbf{x}_0, \mathbf{x}_{i_1}$.

Note that $\mathbf{x}_0 = \lambda_1 \mathbf{x}_{i_1} + (1 - \lambda_1) \mathbf{y}_0$ for some $0 \leq \lambda_1 \leq 1$, that is \mathbf{x}_0 is written as a convex combination of an extreme point \mathbf{x}_{i_1} of \bar{S} and \mathbf{y}_0 which lies on atleast $(k + 1)$ LI hyperplanes defining \bar{S} .

Now repeating the same process by starting with \mathbf{y}_0 , after a finite number of steps we will be able to write $\mathbf{y}_0 = \lambda_2 \mathbf{x}_{i_2} + (1 - \lambda_2) \mathbf{y}_{00}$ for some $0 \leq \lambda_2 \leq 1$, where \mathbf{x}_{i_2} is an extreme point of \bar{S} and \mathbf{y}_{00} lies on atleast $(k + 2)$ LI hyperplanes defining \bar{S} .

Hence $\mathbf{x}_0 = \lambda_1 \mathbf{x}_{i_1} + (1 - \lambda_1) \mathbf{y}_0 = \lambda_1 \mathbf{x}_{i_1} + (1 - \lambda_1)(\lambda_2 \mathbf{x}_{i_2} + (1 - \lambda_2) \mathbf{y}_{00})$
 $= \lambda_1 \mathbf{x}_{i_1} + (1 - \lambda_1) \lambda_2 \mathbf{x}_{i_2} + (1 - \lambda_1)(1 - \lambda_2) \mathbf{y}_{00}$.

That is, $\mathbf{x}_0 = \beta_1 \mathbf{x}_{i_1} + \beta_2 \mathbf{x}_{i_2} + \beta_3 \mathbf{y}_{00}$, where $0 \leq \beta_i \leq 1$ for all $i = 1, 2, 3$ and $\sum_{i=1}^3 \beta_i = 1$.

Continuing this process by starting with \mathbf{y}_{00} , after atmost $n - k - 2$ steps, we will be able to write \mathbf{x}_0 as a convex combination of the extreme points of \bar{S} .

Let $\mathbf{x}_0 = \sum_{j=1}^p \lambda_j \mathbf{x}_{i_j}$, (***)

where $0 \leq \lambda_j \leq 1$ for all $j = 1, 2, \dots, p$ and $\sum_{j=1}^p \lambda_j = 1$.

If all the extreme points in that above expression (of \mathbf{x}_0) are also extreme points of S then we are done.

If not then WLOG let \mathbf{x}_{i_1} be an extreme point of \bar{S} , which is not an extreme point of S .

This implies \mathbf{x}_{i_1} lies on $(n - 1)$ LI defining hyperplanes of S (WLOG assume that the respective normals are $\tilde{\mathbf{a}}_i, i = 1, \dots, n - 1$) and on the added hyperplane H_0 (with a normal $[1, \dots, 1]^T$).

Let $\mathbf{d}_2 \neq \mathbf{0}$ be a vector orthogonal to each of these $(n - 1)$ normals (that is $\tilde{\mathbf{a}}_i^T \mathbf{d}_2 = 0$, for all $i = 1, \dots, n - 1$), (why does this vector exist?) and since $\mathbf{d}_2 \neq \mathbf{0}$, it cannot be orthogonal to the normal of H_0 (why?).

Further $\mathbf{x}_{i_1} \pm \alpha \mathbf{d}_2$, (for any given $\alpha > 0$) cannot both lie on the same closed half space defined

by H_0 and hence cannot both belong to \bar{S} (since \mathbf{x}_{i_1} lies on H_0 that is $[1, \dots, 1]^T \mathbf{x}_{i_1} = M$, and $[1, \dots, 1]^T \mathbf{d}_2 \neq 0$).

WLOG let $\mathbf{x}_{i_1} - \alpha \mathbf{d}_2 \in \bar{S}$.

Since \bar{S} is bounded, there exists $\delta > 0$, sufficiently large such that $\mathbf{x}_{i_1} - \delta \mathbf{d}_2$ is not in \bar{S} .

Let $\theta = \max\{\delta : \mathbf{x}_{i_1} - \delta \mathbf{d}_2 \in \bar{S}\}$ and let $\mathbf{z} = \mathbf{x}_{i_1} - \theta \mathbf{d}_2$, then $\mathbf{z} \in \bar{S}$.

Also \mathbf{z} lies on the $(n-1)$ LI hyperplanes of S on which \mathbf{x}_{i_1} lies and another extra hyperplane of S (it cannot be H_0) which is LI to the the previous $(n-1)$, (since it obstructs indefinite movement along $-\mathbf{d}_2$ starting from \mathbf{x}_{i_1}), hence \mathbf{z} is an extreme point of S .

Check that $\mathbf{z} + \alpha \mathbf{d}_2 \in S$ for all $\alpha > 0$, hence $\mathbf{d}_2 \neq \mathbf{0}$ is a direction of S . Since \mathbf{d}_2 satisfies $\tilde{\mathbf{a}}_i^T \mathbf{d} = 0$, for all $i = 1, \dots, n-1$, that is it lies on $(n-1)$ LI hyperplanes defining D (the set of directions of S), hence \mathbf{d}_2 is an extreme direction of S .

Also since $\mathbf{x}_{i_1} = \mathbf{z} + \theta \mathbf{d}_2$, if we substitute this expression of \mathbf{x}_{i_1} in (**), and do this similarly for all other extreme points of \bar{S} which are not extreme points of S in (***) then finally we would have written \mathbf{x}_0 as a linear combination of the extreme points of S plus a nonnegative linear combination of the extreme directions of S .

Remark: If $S \neq \emptyset$ is bounded then there is no need to add H_0 to the the existing set of $(m+n)$ defining hyperplanes of S in the above proof, and the process followed above terminates at (**).

Observation 6: If $S = \text{Fea}(LPP)$ is a nonempty bounded set then any $\mathbf{x} \in S$ can be written as a convex combination of the extreme points of S .

Observation 7: Since D' , the set of distinct directions of S (if it is nonempty) is a bounded set because of the constraints $\mathbf{d} \geq \mathbf{0}$ and $\sum_{i=1}^n d_i = 1$, so any $\mathbf{d} \in D'$ can be written as a convex combination of the extreme points of D' . So any direction $\mathbf{d} \in D$ of S can be written as a nonnegative linear combination of the extreme directions of S .

Observation 8: Note that if there exists a $\mathbf{d} \in D$ such that $\mathbf{c}^T \mathbf{d} < 0$ then the LPP(*)
(*) Min $\mathbf{c}^T \mathbf{x}$, subject to $A\mathbf{x} \leq \mathbf{b}$, $\mathbf{x} \geq \mathbf{0}$)

does not have an optimal solution.

Since for any given $\mathbf{x} \in S$, $\mathbf{x} + \alpha \mathbf{d} \in S$ for all $\alpha \geq 0$ and

$\mathbf{c}^T (\mathbf{x} + \alpha \mathbf{d}) = \mathbf{c}^T \mathbf{x} + \alpha \mathbf{c}^T \mathbf{d}$ can be made smaller than any real M , by choosing $\alpha > 0$ sufficiently large.

Exercise: If $\mathbf{c}^T \mathbf{d}_j \geq 0$ for all extreme directions \mathbf{d}_j of the nonempty and unbounded feasible region S of a LPP, then does it imply that $\mathbf{c}^T \mathbf{d} \geq 0$ for all directions $\mathbf{d} \in D$, of the feasible region S ?

Ans is **yes**, since any $\mathbf{d} \in D$ can be written as a **nonnegative** linear combinations of the extreme directions of S , that is,

$\mathbf{d} = \sum_{j=1}^r \mu_j \mathbf{d}_j$, for some $\mu_j \geq 0$ for all $j = 1, 2, \dots, r$,

where \mathbf{d}_j 's are the (instead of writing **the**, should be more correctly written as, a set of) extreme directions of S .

Hence $\mathbf{c}^T \mathbf{d} = \sum_{j=1}^r \mu_j \mathbf{c}^T \mathbf{d}_j \geq 0$.

Observation 9: From the representation theorem of S we can see that if $S \neq \emptyset$, $D \neq \emptyset$ and $\mathbf{c}^T \mathbf{d}_j \geq 0$ for all $j = 1, 2, \dots, r$, then LPP(*) has an optimal solution, and atleast one of the optimal solutions is an extreme point of S .

If $\mathbf{c}^T \mathbf{d}_j \geq 0$ for all $j = 1, 2, \dots, r$, then for all $\mathbf{x} \in S$,

$$\mathbf{c}^T \mathbf{x} = \sum_{i=1}^k \lambda_i \mathbf{c}^T \mathbf{x}_i + \sum_{j=1}^r \mu_j \mathbf{c}^T \mathbf{d}_j \geq \sum_{i=1}^k \lambda_i \mathbf{c}^T \mathbf{x}_i \quad (1)$$

where $0 \leq \lambda_i \leq 1$ for all $i = 1, 2, \dots, k$, $\sum_{i=1}^k \lambda_i = 1$, and $\mu_j \geq 0$, for all $j = 1, 2, \dots, r$.

If \mathbf{x}_{i_0} is the the extreme point such that,

$\mathbf{c}^T \mathbf{x}_{i_0} = \min\{\mathbf{c}^T \mathbf{x}_i : i = 1, 2, \dots, k\}$, (note that $i_0 \in \{1, 2, \dots, k\}$) then from (1),

$\mathbf{c}^T \mathbf{x} \geq \sum_{i=1}^k \lambda_i \mathbf{c}^T \mathbf{x}_i \geq (\sum_{i=1}^k \lambda_i) \mathbf{c}^T \mathbf{x}_{i_0} = \mathbf{c}^T \mathbf{x}_{i_0}$, for all $\mathbf{x} \in S$.

Hence the LPP(*) has an optimal solution, and the extreme point \mathbf{x}_{i_0} of S is an optimal solution.

Observation 10: From the representation theorem of S we can also see that if $S = \text{Fea}(LPP)$ is nonempty and bounded then

(1) the LPP(*) has an optimal solution

(2) and atleast one of the extreme points is an optimal solution.

If S is bounded then for all $\mathbf{x} \in S$, $\mathbf{c}^T \mathbf{x} = \sum_{i=1}^k \lambda_i \mathbf{c}^T \mathbf{x}_i$ for some λ_i , $i = 1, \dots, k$ where $0 \leq \lambda_i \leq 1$ for all $i = 1, 2, \dots, k$, $\sum_{i=1}^k \lambda_i = 1$.

Again take \mathbf{x}_{i_0} as the the extreme point such that,

$\mathbf{c}^T \mathbf{x}_{i_0} = \min\{\mathbf{c}^T \mathbf{x}_i : i = 1, 2, \dots, k\}$,

then by repeating the calculations of the previous proof we get $\mathbf{c}^T \mathbf{x} \geq \mathbf{c}^T \mathbf{x}_{i_0}$ for all $\mathbf{x} \in S$.

Hence the LPP(*) has an optimal solution

(we already know one proof of this by Weierstrass the above gives an alternate proof)

and the extreme point \mathbf{x}_{i_0} is an optimal solution.

From the above observations we can conclude the following:

Conclusion 1: If $S = \text{Fea}(LPP) \neq \emptyset$, then the LPP (*) has an optimal solution if and only if one of the following is true:

(i) $S = \text{Fea}(LPP)$ is bounded (also seen before by using extreme value theorem)

(ii) $S = \text{Fea}(LPP)$ is unbounded and $\mathbf{c}^T \mathbf{d}_j \geq 0$ for all extreme directions \mathbf{d}_j of the feasible region S (follows from observation 6 and observation 7).

Conclusion 2: If LPP (*) has an optimal solution then there exists an extreme point of the feasible region S , which is an optimal solution.

Exercise: Give an example of a **nonlinear** function $f : S \rightarrow \mathbb{R}$, where $S \subset \mathbb{R}$ is a closed and bounded polyhedral subset of \mathbb{R} , (what are these sets?) such that f has a minimum value in S but the minimum value is not attained at any extreme point of S .

Conclusion 3: If $S = \text{Fea}(LPP)$ is nonempty, and there exists an $M \in \mathbb{R}$ such that for all $\mathbf{x} \in S$, $\mathbf{c}^T \mathbf{x} \geq M$, then the LPP (*) has an optimal solution.

To understand the significance of the previous result solve the following problems.

Exercise: Give an example of a **linear** function $f : S \rightarrow \mathbb{R}$, where $S \subset \mathbb{R}$ is not a polyhedral subset of \mathbb{R} , such that $f(x) \geq 1$ but f does not have a minimum value in S .

Exercise: Give an example of a **nonlinear** function $f : S \rightarrow \mathbb{R}$, where $S \subset \mathbb{R}$ is a polyhedral subset of \mathbb{R} , such that $f(x) \geq 1$ but f does not have a minimum value in S .

We can come to similar conclusions if we consider a linear programming problem, LPP(**) as

(**) Max $\mathbf{c}^T \mathbf{x}$
subject to $A\mathbf{x} \leq \mathbf{b}$, $\mathbf{x} \geq \mathbf{0}$.

Conclusion 1a: If $S = \text{Fea}(LPP) \neq \emptyset$, then the LPP (**) has an optimal solution if and only if one of the following is true:

- (i) $S = \text{Fea}(LPP)$ is bounded
- (ii) $S = \text{Fea}(LPP)$ is unbounded and $\mathbf{c}^T \mathbf{d}_j \leq 0$ for all extreme directions \mathbf{d}_j of the feasible region S .

Conclusion 2a: If a LPP (**) has an optimal solution then there exists an extreme point of the feasible region S , which is an optimal solution.

Conclusion 3a: If $S = \text{Fea}(LPP)$ is nonempty, and there exists an $M \in \mathbb{R}$ such that for all $\mathbf{x} \in S$, $\mathbf{c}^T \mathbf{x} \leq M$, then the LPP (**) has an optimal solution.

Appendix: (Optional reading) To show that if the feasible region of a linear programming problem is unbounded then it should have atleast one direction.

Proof: Let us assume representation theorem to be true for all nonempty bounded feasible regions. Let S have extreme points $\mathbf{x}_1, \dots, \mathbf{x}_r$ and consider $S \cap H$, where $H = \{\mathbf{x} \in \mathbb{R}^n : \sum_{i=1}^n x_i \leq M\}$ and $M > 0$ is such that $\mathbf{x}_1, \dots, \mathbf{x}_r \in S \cap H$ but none of $\mathbf{x}_1, \dots, \mathbf{x}_r$ lie on the hyperplane corresponding to H .

Let $\mathbf{x}_0 \in S$ be such that $M \geq \sum_{i=1}^n x_{0i} > \max\{\sum_{i=1}^n x_{ki} : k = 1, \dots, r\}$ (there exists such an \mathbf{x}_0 , why?).

Check that this \mathbf{x}_0 **cannot** be written as a convex combination of the extreme points of S (check this), but it can be written as a convex combination of the extreme points of $S \cap H$, at least one of which (call that extreme point \mathbf{u}) must lie on the hyperplane corresponding to H .

Then by repeating the proof of representation theorem (for unbounded feasible region) conclude that $\mathbf{u} = \mathbf{x}_i + \alpha \mathbf{d}$ for some $\alpha > 0$ and \mathbf{x}_i an extreme point of S and \mathbf{d} an extreme direction of S where \mathbf{d} is orthogonal to the normals of all the hyperplanes on which \mathbf{u} lies except the normal corresponding to H .