Foundations of the Simplex Method

K. Subramani[1]

[1] Lane Department of Computer Science and Electrical Engineering
West Virginia University

January 4, 2013

Outline

1. Graphical solutions to two dimensional problems
   - Representing constraints as sections of the plane
   - Handling the objective function
   - Exercises

Outline

Outline

Outline

Outline

Geometrically representing linear programs

In addition to having a standard form, it is also helpful to understand systems of constraints geometrically. To gain such understanding we will learn how to solve two variable linear programs geometrically.

Geometrically representing linear programs

In addition to having a standard form, it is also helpful to understand systems of constraints geometrically. To gain such understanding we will learn how to solve two variable linear programs geometrically.

Geometric View of Constraints

First we will see how constrains can be considered to be portions of the x_1,x_2 plane.

Geometrically representing linear programs

In addition to having a standard form, it is also helpful to understand systems of constraints geometrically. To gain such understanding we will learn how to solve two variable linear programs geometrically.

Geometric View of Constraints

First we will see how constrains can be considered to be portions of the x_1,x_2 plane. An equality, such as $x_1 + x_2 = 3$, can be viewed as a line in the x_1,x_2 plane.

Geometrically representing linear programs

In addition to having a standard form, it is also helpful to understand systems of constraints geometrically. To gain such understanding we will learn how to solve two variable linear programs geometrically.

Geometric View of Constraints

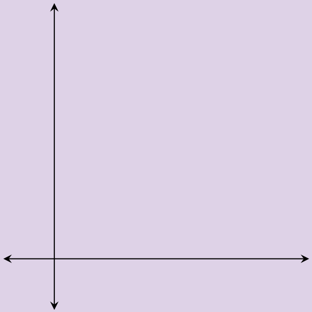
First we will see how constrains can be considered to be portions of the x_1,x_2 plane. An equality, such as $x_1 + x_2 = 3$, can be viewed as a line in the x_1,x_2 plane.

Geometrically representing linear programs

In addition to having a standard form, it is also helpful to understand systems of constraints geometrically. To gain such understanding we will learn how to solve two variable linear programs geometrically.

Geometric View of Constraints

First we will see how constrains can be considered to be portions of the x_1,x_2 plane. An equality, such as $x_1 + x_2 = 3$, can be viewed as a line in the x_1,x_2 plane.

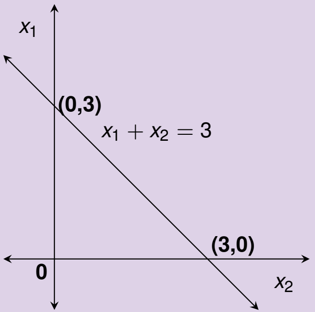


Geometrically representing linear programs

In addition to having a standard form, it is also helpful to understand systems of constraints geometrically. To gain such understanding we will learn how to solve two variable linear programs geometrically.

Geometric View of Constraints

First we will see how constrains can be considered to be portions of the x_1,x_2 plane. An equality, such as $x_1 + x_2 = 3$, can be viewed as a line in the x_1,x_2 plane.



Geometric View of Constraints

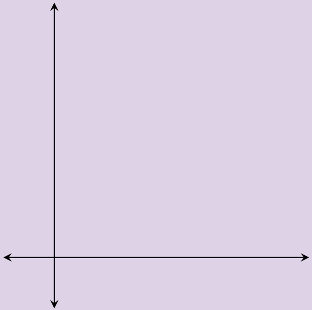
Similarly an inequality, such as $x_1 + x_2 \leq 3$, can be viewed as as the half plane above or below a line in the x_1,x_2 plane.

Geometric View of Constraints

Similarly an inequality, such as $x_1 + x_2 \leq 3$, can be viewed as as the half plane above or below a line in the x_1,x_2 plane.

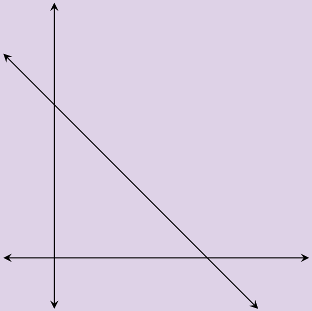
Geometric View of Constraints

Similarly an inequality, such as $x_1 + x_2 \leq 3$, can be viewed as as the half plane above or below a line in the x_1,x_2 plane.



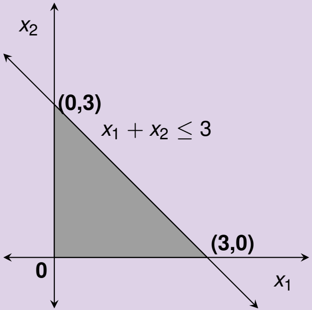
Geometric View of Constraints

Similarly an inequality, such as $x_1 + x_2 \leq 3$, can be viewed as as the half plane above or below a line in the x_1,x_2 plane.



Geometric View of Constraints

Similarly an inequality, such as $x_1 + x_2 \leq 3$, can be viewed as as the half plane above or below a line in the x_1,x_2 plane.

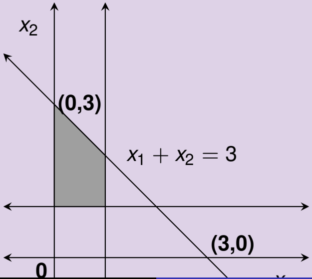


Geometric View of Constraints

For a system of constraints the section of the plane corresponding to solutions to that system is simply the intersection of the portions of the plane corresponding to each constraint. For instance, the constraints

$$\begin{aligned} x_1 &\leq 1 \\ x_2 &\geq 1 \\ x_1 + x_2 &\leq 3 \\ x_1, x_2 &\geq 0 \end{aligned}$$

would produce



Outline

Geometric representation of the objective function

Now that we know how to represent constraints geometrically, we need to consider the objective function.

Geometric representation of the objective function

Now that we know how to represent constraints geometrically, we need to consider the objective function. For a fixed z, the objective function is simply an equality, and can thus be represented as a line in the x_1, x_2 plane.

Geometric representation of the objective function

Now that we know how to represent constraints geometrically, we need to consider the objective function. For a fixed z, the objective function is simply an equality, and can thus be represented as a line in the x_1,x_2 plane. If we allow z to vary then the objective function can be represented as a series of parallel lines each corresponding to a different value for z.

Geometric representation of the objective function

Now that we know how to represent constraints geometrically, we need to consider the objective function. For a fixed z, the objective function is simply an equality, and can thus be represented as a line in the x_1,x_2 plane. If we allow z to vary then the objective function can be represented as a series of parallel lines each corresponding to a different value for z.
If we are trying to maximize z then we find the maximum z for which the corresponding line passes though the portion of the plane corresponding to the system of constraints.

Geometric representation of the objective function

Now that we know how to represent constraints geometrically, we need to consider the objective function. For a fixed z, the objective function is simply an equality, and can thus be represented as a line in the x_1,x_2 plane. If we allow z to vary then the objective function can be represented as a series of parallel lines each corresponding to a different value for z.

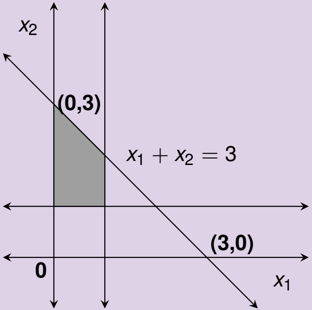
If we are trying to maximize z then we find the maximum z for which the corresponding line passes though the portion of the plane corresponding to the system of constraints. When minimizing we consider the minimum such z.

Geometric representation of the objective function

Now that we know how to represent constraints geometrically, we need to consider the objective function. For a fixed z, the objective function is simply an equality, and can thus be represented as a line in the x_1,x_2 plane. If we allow z to vary then the objective function can be represented as a series of parallel lines each corresponding to a different value for z.
If we are trying to maximize z then we find the maximum z for which the corresponding line passes though the portion of the plane corresponding to the system of constraints. When minimizing we consider the minimum such z.
It also helps to find the gradient of z as it identifies the direction in which z grows the fastest. As the objective function is of the form $z = c_1 x_1 + c_2 x_2$, the gradient is simply the vector (c_1, c_2).

Handling the objective function

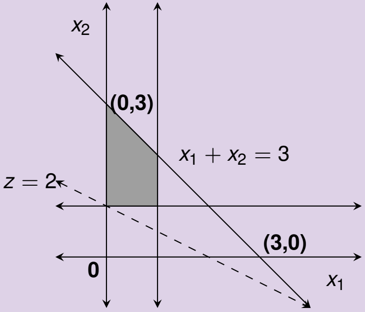
For example adding the objective function $z = x_1 + 2x_2$ to our previous example yields



Thus trying to maximize z would yield that maximum z to be 6 when $x_1 = 0$ and $x_2 = 3$. Similarly trying to minimize z would yield that minimum z to be 2 when $x_1 = 0$ and $x_2 = 1$.

Handling the objective function

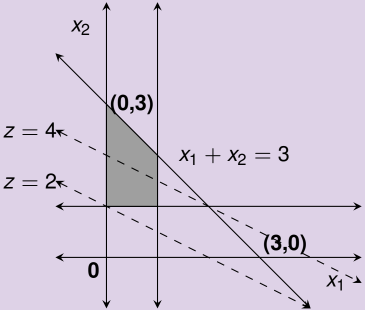
For example adding the objective function $z = x_1 + 2x_2$ to our previous example yields



Thus trying to maximize z would yield that maximum z to be 6 when $x_1 = 0$ and $x_2 = 3$. Similarly trying to minimize z would yield that minimum z to be 2 when $x_1 = 0$ and $x_2 = 1$.

Handling the objective function

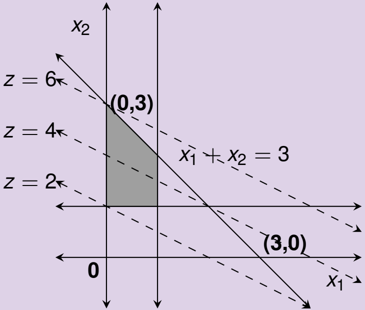
For example adding the objective function $z = x_1 + 2x_2$ to our previous example yields



Thus trying to maximize z would yield that maximum z to be 6 when $x_1 = 0$ and $x_2 = 3$. Similarly trying to minimize z would yield that minimum z to be 2 when $x_1 = 0$ and $x_2 = 1$.

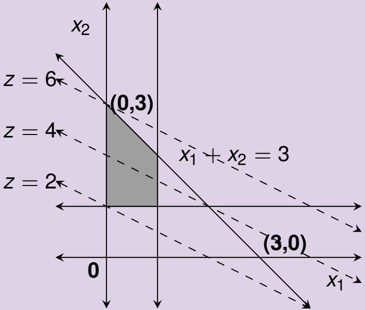
Handling the objective function

For example adding the objective function $z = x_1 + 2x_2$ to our previous example yields



Handling the objective function

For example adding the objective function $z = x_1 + 2x_2$ to our previous example yields



Thus trying to maximize z would yield that maximum z to be 6 when $x_1 = 0$ and $x_2 = 3$. Similarly trying to minimize z would yield that minimum z to be 2 when $x_1 = 0$ and $x_2 = 1$.

Outline

Exercise 1

Solve the following linear program graphically

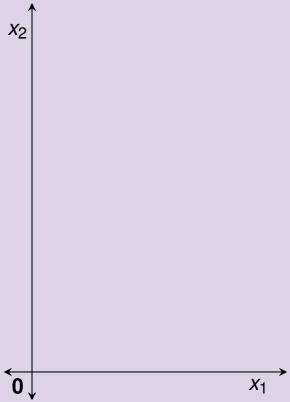
$$\text{minimize } z = 4x_1 + 5x_2$$

subject to

$$\begin{aligned} 3x_1 + 2x_2 &\leq 24 \\ x_1 &\geq 5 \\ 3x_1 - x_2 &\leq 6 \\ x_1, x_2 &\geq 0 \end{aligned}$$

Solution

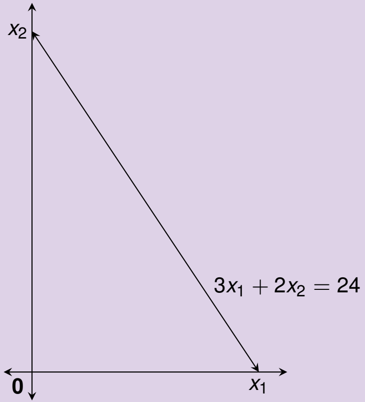
If the constraints are plotted onto a graph we see



There are no points which satisfy all three constraints. Thus no solution exists.

Solution

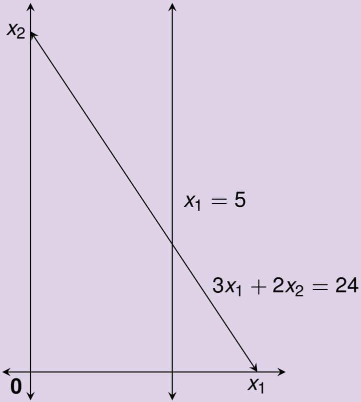
If the constraints are plotted onto a graph we see



There are no points which satisfy all three constraints. Thus no solution exists.

Solution

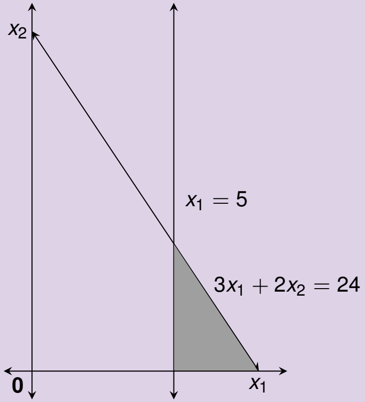
If the constraints are plotted onto a graph we see



There are no points which satisfy all three constraints. Thus no solution exists.

Solution

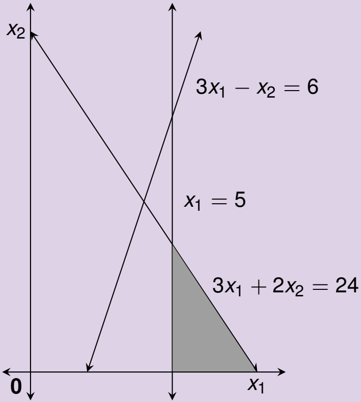
If the constraints are plotted onto a graph we see



There are no points which satisfy all three constraints. Thus no solution exists.

Solution

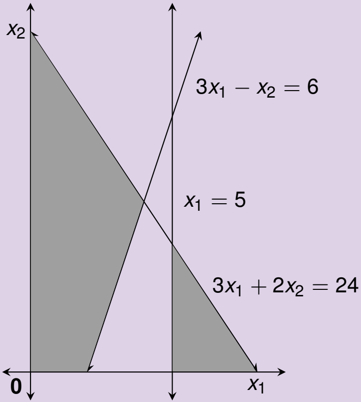
If the constraints are plotted onto a graph we see



There are no points which satisfy all three constraints. Thus no solution exists.

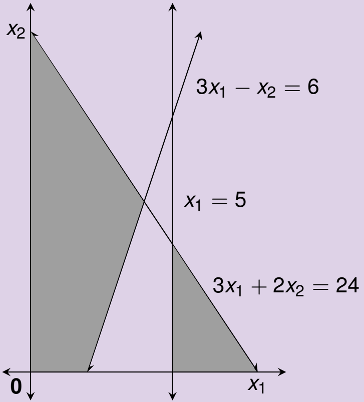
Solution

If the constraints are plotted onto a graph we see



Solution

If the constraints are plotted onto a graph we see



There are no points which satisfy all three constraints. Thus no solution exists.

Exercise 2

Solve the following system of constraints graphically

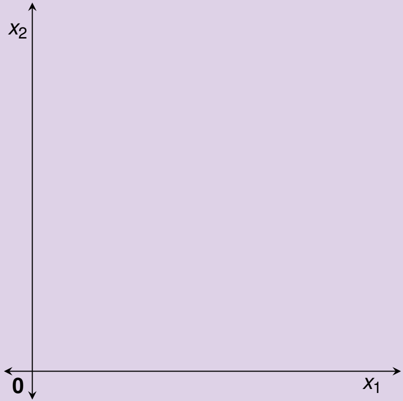
$$\text{minimize } z = x_1 - 4x_2$$

subject to

$$\begin{array}{rcl} x_1 + x_2 & \leq & 12 \\ -2x_1 + x_2 & \leq & 4 \\ x_2 & \leq & 8 \\ x_1 - 3x_2 & \leq & 4 \\ x_1, x_2 & \geq & 0 \end{array}$$

Solution

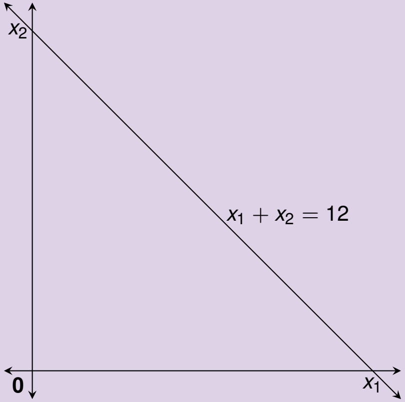
Plotting the constraints and then checking various values of z we get.



Thus the minimum value of z is $z = -30$ and occurs at $(x_1, x_2) = (2, 8)$.

Solution

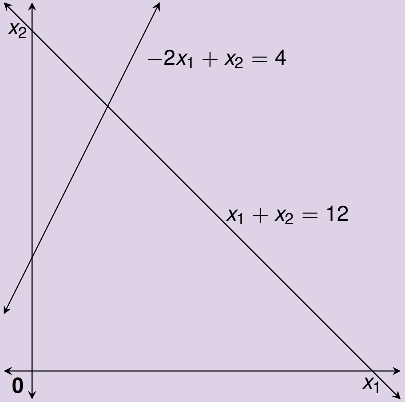
Plotting the constraints and then checking various values of z we get.



Thus the minimum value of z is $z = -30$ and occurs at $(x_1, x_2) = (2, 8)$.

Solution

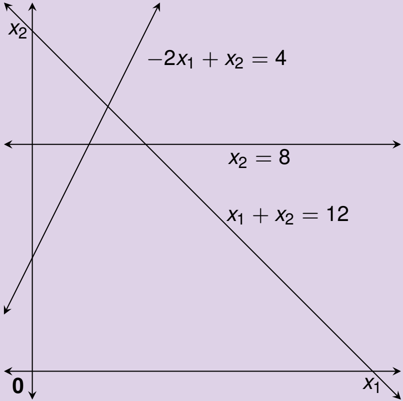
Plotting the constraints and then checking various values of z we get.



Thus the minimum value of z is $z = -30$ and occurs at $(x_1, x_2) = (2, 8)$.

Solution

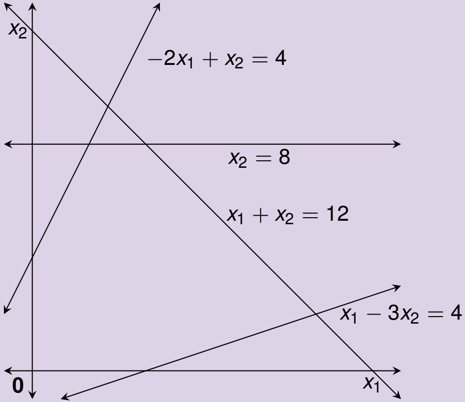
Plotting the constraints and then checking various values of z we get.



Thus the minimum value of z is $z = -30$ and occurs at $(x_1, x_2) = (2, 8)$.

Solution

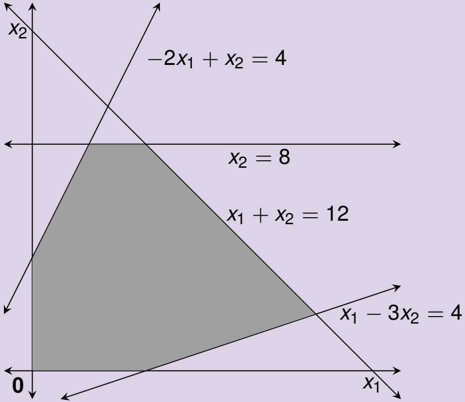
Plotting the constraints and then checking various values of z we get.



Thus the minimum value of z is $z = -30$ and occurs at $(x_1, x_2) = (2, 8)$.

Solution

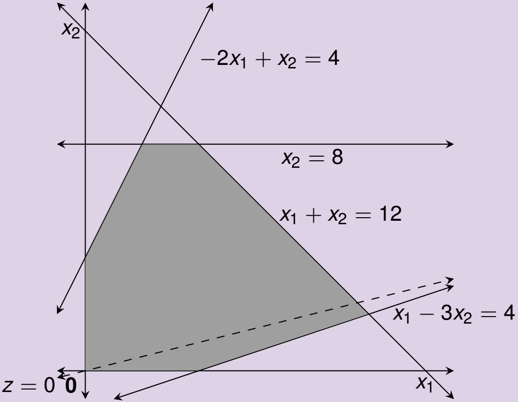
Plotting the constraints and then checking various values of z we get.



Thus the minimum value of z is $z = -30$ and occurs at $(x_1, x_2) = (2, 8)$.

Solution

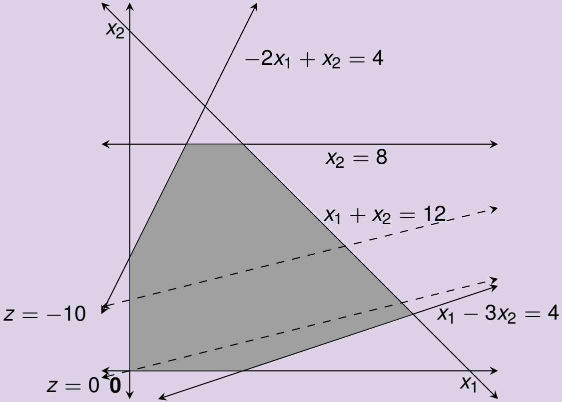
Plotting the constraints and then checking various values of z we get.



Thus the minimum value of z is $z = -30$ and occurs at $(x_1, x_2) = (2, 8)$.

Solution

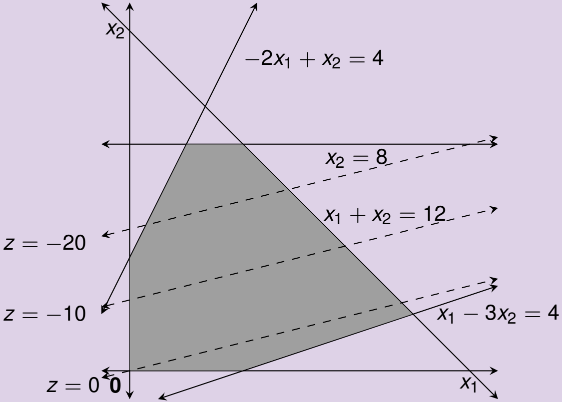
Plotting the constraints and then checking various values of z we get.



Thus the minimum value of z is $z = -30$ and occurs at $(x_1, x_2) = (2, 8)$.

Solution

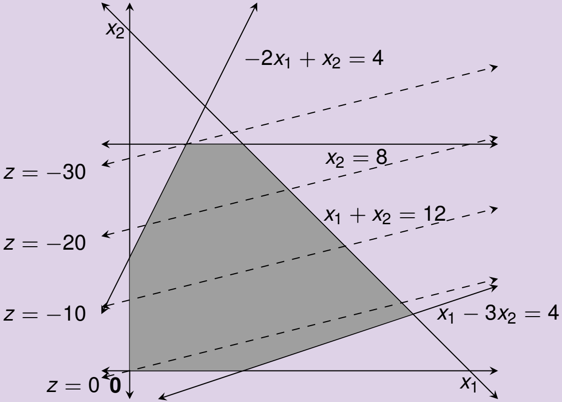
Plotting the constraints and then checking various values of z we get.



Thus the minimum value of z is $z = -30$ and occurs at $(x_1, x_2) = (2, 8)$.

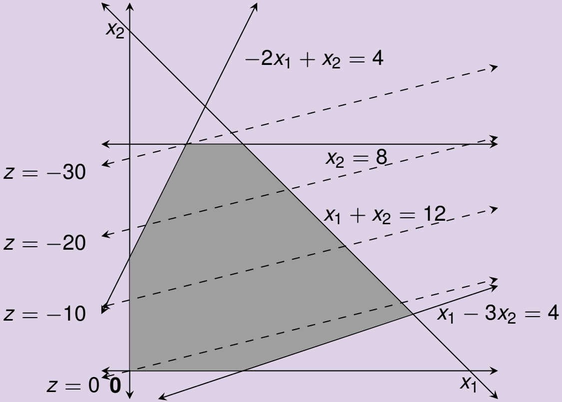
Solution

Plotting the constraints and then checking various values of z we get.



Solution

Plotting the constraints and then checking various values of z we get.



Thus the minimum value of z is $z = -30$ and occurs at $(x_1, x_2) = (2, 8)$.

Exercise 3

Solve the following linear program graphically

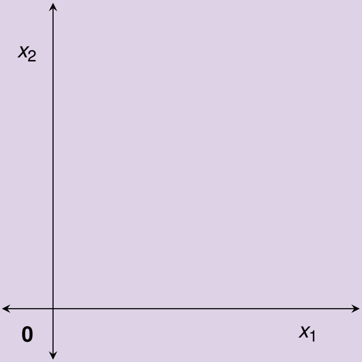
$$\text{maximize } z = x_1 + 2x_2$$

subject to

$$\begin{aligned}
-2x_1 + x_2 &\leq 2 \\
2x_1 + 5x_2 &\geq 10 \\
x_1 - 4x_2 &\leq 2 \\
x_1, x_2 &\geq 0
\end{aligned}$$

Solution

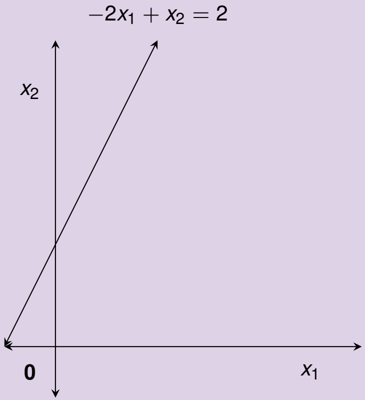
Plotting the constraints and then checking various values of z we get.



Thus there is no maximum value of z as z can be increased indefinitely and the system will still be feasible.

Solution

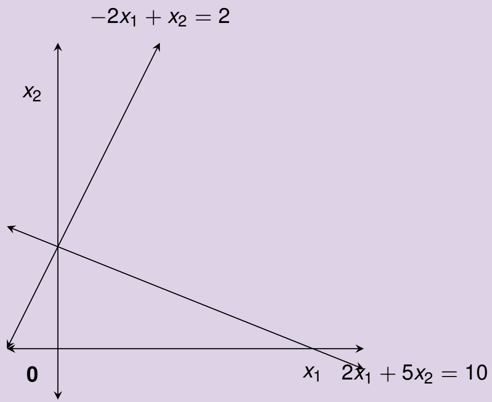
Plotting the constraints and then checking various values of z we get.



Thus there is no maximum value of z as z can be increased indefinitely and the system will still be feasible.

Solution

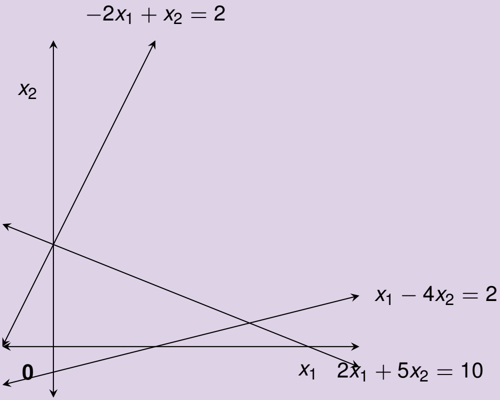
Plotting the constraints and then checking various values of z we get.



Thus there is no maximum value of z as z can be increased indefinitely and the system will still be feasible.

Solution

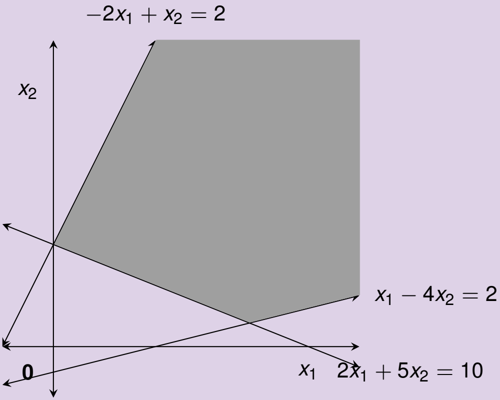
Plotting the constraints and then checking various values of z we get.



Thus there is no maximum value of z as z can be increased indefinitely and the system will still be feasible.

Solution

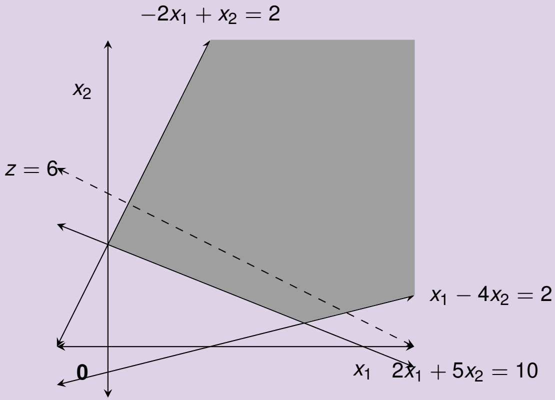
Plotting the constraints and then checking various values of z we get.



Thus there is no maximum value of z as z can be increased indefinitely and the system will still be feasible.

Solution

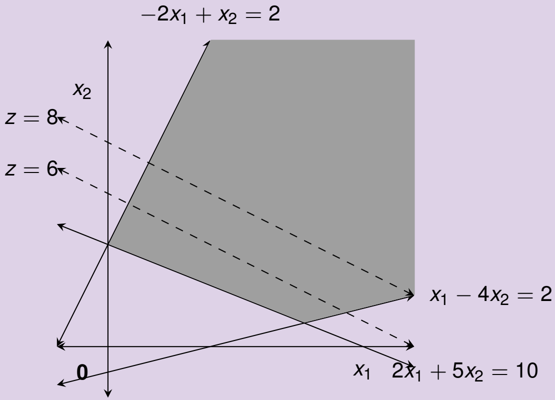
Plotting the constraints and then checking various values of z we get.



Thus there is no maximum value of z as z can be increased indefinitely and the system will still be feasible.

Solution

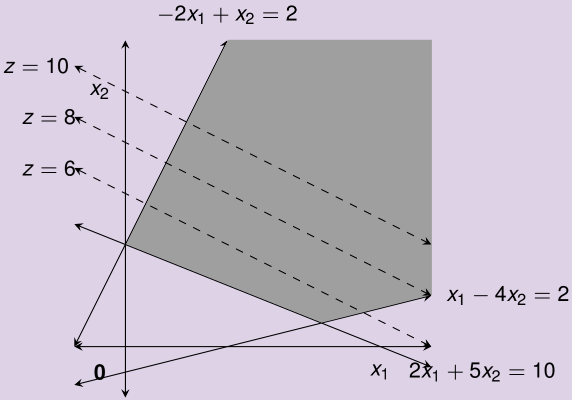
Plotting the constraints and then checking various values of z we get.



Thus there is no maximum value of z as z can be increased indefinitely and the system will still be feasible.

Solution

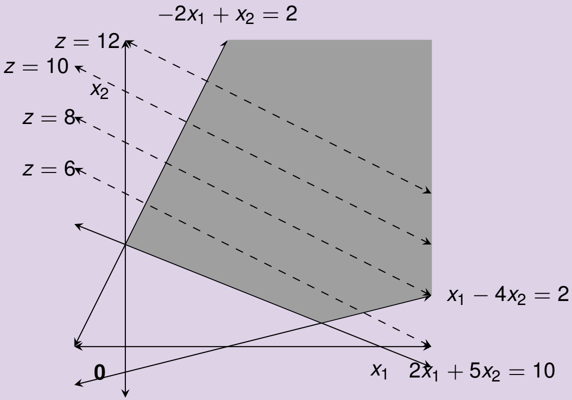
Plotting the constraints and then checking various values of z we get.



Thus there is no maximum value of z as z can be increased indefinitely and the system will still be feasible.

Solution

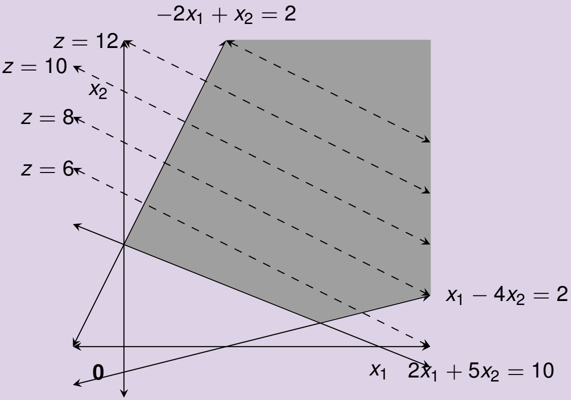
Plotting the constraints and then checking various values of z we get.



Thus there is no maximum value of z as z can be increased indefinitely and the system will still be feasible.

Solution

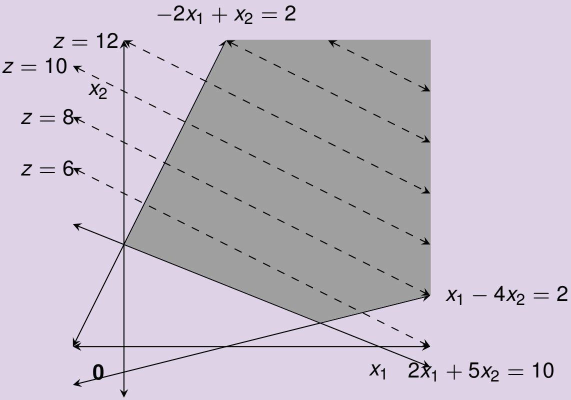
Plotting the constraints and then checking various values of z we get.



Thus there is no maximum value of z as z can be increased indefinitely and the system will still be feasible.

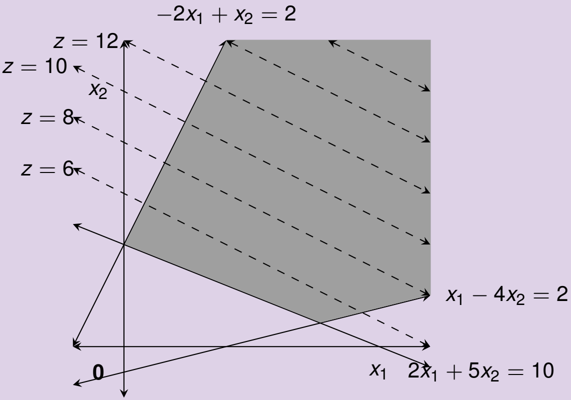
Solution

Plotting the constraints and then checking various values of z we get.



Solution

Plotting the constraints and then checking various values of z we get.



Thus there is no maximum value of z as z can be increased indefinitely and the system will still be feasible.

Outline

Definition (Hyperplane)

A hyperplane is a set of points, $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$, that satisfy $\mathbf{ax} = b$, where $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and b is a scalar.

Definition (Hyperplane)

A hyperplane is a set of points, $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$, that satisfy $\mathbf{ax} = b$, where $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and b is a scalar.

Definition (Halfspace)

A closed halfspace corresponding to a hyperplane $\mathbf{a} \cdot \mathbf{x} = b$ is either of the sets $H^+ = \{\mathbf{x} : \mathbf{a} \cdot \mathbf{x} \geq b\}$ or $H^- = \{\mathbf{x} : \mathbf{a} \cdot \mathbf{x} \leq b\}$. If the inequalities involved are strict then the corresponding halfspace are referred to as open halfspaces.

Outline

Definition (Convex Set)

A set, S, is convex if for any two points $\mathbf{x}_1, \mathbf{x}_2 \in S$ then all points on the line segment connecting $\mathbf{x}_1$ and $\mathbf{x}_2$ are in S. This means that $\forall \alpha \in [0, 1]$, $\alpha \cdot \mathbf{x}_1 + (1 - \alpha) \cdot \mathbf{x}_2 \in S$.

Definition (Convex Set)

A set, S, is convex if for any two points $\mathbf{x}_1, \mathbf{x}_2 \in S$ then all points on the line segment connecting $\mathbf{x}_1$ and $\mathbf{x}_2$ are in S. This means that $\forall \alpha \in [0, 1]$, $\alpha \cdot \mathbf{x}_1 + (1 - \alpha) \cdot \mathbf{x}_2 \in S$.

Definition (Polyhedral Set)

A set S is polyhedral if it is the intersection of a finite number of halfspaces.

Definition (Convex Set)

A set, S, is convex if for any two points $\mathbf{x}_1, \mathbf{x}_2 \in S$ then all points on the line segment connecting $\mathbf{x}_1$ and $\mathbf{x}_2$ are in S. This means that $\forall \alpha \in [0, 1]$, $\alpha \cdot \mathbf{x}_1 + (1 - \alpha) \cdot \mathbf{x}_2 \in S$.

Definition (Polyhedral Set)

A set S is polyhedral if it is the intersection of a finite number of halfspaces.

Systems of constraints as Polyhedral Sets

A constraint system of the form $S = \{\mathbf{x} : \mathbf{A} \cdot \mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ is a polyhedral set as each constraint corresponds to a halfspace.

Theorem

The set $S = \{\mathbf{x} : \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ *is convex.*

Theorem

The set $S = \{\mathbf{x} : \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ *is convex.*

Proof.

Let $\mathbf{x}_1, \mathbf{x}_2 \in S$, and let $\alpha \in [0, 1]$.

Theorem

The set $S = \{\mathbf{x} : \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ *is convex.*

Proof.

Let $\mathbf{x}_1, \mathbf{x}_2 \in S$, and let $\alpha \in [0, 1]$. So, by definition of S, $\mathbf{A} \cdot \mathbf{x}_1 = \mathbf{b}$ and $\mathbf{A} \cdot \mathbf{x}_2 = \mathbf{b}$.
Thus for any $\alpha \in [0, 1]$, we have that
$\mathbf{A} \cdot (\alpha \cdot \mathbf{x}_1 + (1 - \alpha) \cdot \mathbf{x}_2) = \alpha \cdot \mathbf{A} \cdot \mathbf{x}_1 + (1 - \alpha) \cdot \mathbf{A} \cdot \mathbf{x}_2 = \alpha \cdot \mathbf{b} + (1 - \alpha) \cdot \mathbf{b} = \mathbf{b}$.

Theorem

The set $S = \{\mathbf{x} : \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ *is convex.*

Proof.

Let $\mathbf{x}_1, \mathbf{x}_2 \in S$, and let $\alpha \in [0, 1]$. So, by definition of S, $\mathbf{A} \cdot \mathbf{x}_1 = \mathbf{b}$ and $\mathbf{A} \cdot \mathbf{x}_2 = \mathbf{b}$. Thus for any $\alpha \in [0, 1]$, we have that $\mathbf{A} \cdot (\alpha \cdot \mathbf{x}_1 + (1 - \alpha) \cdot \mathbf{x}_2) = \alpha \cdot \mathbf{A} \cdot \mathbf{x}_1 + (1 - \alpha) \cdot \mathbf{A} \cdot \mathbf{x}_2 = \alpha \cdot \mathbf{b} + (1 - \alpha) \cdot \mathbf{b} = \mathbf{b}$. And as $\alpha, (1 - \alpha) \geq 0$ and $\mathbf{x}_1, \mathbf{x}_2 \geq \mathbf{0}$ we have that $\alpha \cdot \mathbf{x}_1 + (1 - \alpha) \cdot \mathbf{x}_2 \geq \mathbf{0}$. Thus for any $\alpha \in [0, 1]$ $\alpha \cdot \mathbf{x}_1 + (1 - \alpha) \cdot \mathbf{x}_2 \in S$. Thus S is convex. $\square$

Outline

Definition (Extreme Point)

A point **x** in a convex set S is said to be an extreme point if it does not lie on the interior of a line segment connecting two distinct points in S.

Definition (Extreme Point)

A point $\mathbf{x}$ in a convex set S is said to be an extreme point if it does not lie on the interior of a line segment connecting two distinct points in S. Mathematically there do not exist $\mathbf{x}_1, \mathbf{x}_2 \in S$, $\mathbf{x}_1 \neq \mathbf{x}_2$, and $\alpha \in (0, 1)$ such that $\mathbf{x} = \alpha \cdot \mathbf{x}_1 + (1 - \alpha) \cdot \mathbf{x}_2$.

Definition (Extreme Point)

A point $\mathbf{x}$ in a convex set S is said to be an extreme point if it does not lie on the interior of a line segment connecting two distinct points in S. Mathematically there do not exist $\mathbf{x}_1, \mathbf{x}_2 \in S$, $\mathbf{x}_1 \neq \mathbf{x}_2$, and $\alpha \in (0,1)$ such that $\mathbf{x} = \alpha \cdot \mathbf{x}_1 + (1-\alpha) \cdot \mathbf{x}_2$.

Definition (Direction)

A non-zero vector $\mathbf{d} = (d_1, d_2, \ldots, d_n)$ is a direction of a convex set S if $\forall \mathbf{x} \in S$ and $\forall \lambda \geq 0$, $\mathbf{x} + \lambda \cdot \mathbf{d} \in S$.

Definition (Extreme Point)

A point $\mathbf{x}$ in a convex set S is said to be an extreme point if it does not lie on the interior of a line segment connecting two distinct points in S. Mathematically there do not exist $\mathbf{x}_1, \mathbf{x}_2 \in S$, $\mathbf{x}_1 \neq \mathbf{x}_2$, and $\alpha \in (0, 1)$ such that $\mathbf{x} = \alpha \cdot \mathbf{x}_1 + (1 - \alpha) \cdot \mathbf{x}_2$.

Definition (Direction)

A non-zero vector $\mathbf{d} = (d_1, d_2, \ldots, d_n)$ is a direction of a convex set S if $\forall \mathbf{x} \in S$ and $\forall \lambda \geq 0$, $\mathbf{x} + \lambda \cdot \mathbf{d} \in S$.

Definition (Extreme Direction)

A direction $\mathbf{d}$ of a convex set S is said to be an extreme direction if cannot be expressed as a positive combination of two distinct directions of S.

Outline

Goal

First we want to develop a method of identifying the extreme points of a system of constraints in standard form.

Theorem

Let $S = \{\mathbf{x} : \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, where $\mathbf{A}$ is $m \times n$ and rank$(\mathbf{A}) = m < n$. $\bar{\mathbf{x}}$ is an extreme point of S if and only if $\bar{\mathbf{x}}$ is the intersection of n linearly independent hyperplanes.

Goal

First we want to develop a method of identifying the extreme points of a system of constraints in standard form.

Theorem

Let $S = \{\mathbf{x} : \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, *where* $\mathbf{A}$ *is* $m \times n$ *and* $\text{rank}(\mathbf{A}) = m < n$. $\bar{\mathbf{x}}$ *is an extreme point of* S *if and only if* $\bar{\mathbf{x}}$ *is the intersection of* n *linearly independent hyperplanes.*

only if.

Let $\bar{\mathbf{x}}$ be an extreme point of S. To get a contradiction we will assume that $\bar{\mathbf{x}}$ lies on less than n linearly independent hyperplanes.

Goal

First we want to develop a method of identifying the extreme points of a system of constraints in standard form.

Theorem

Let $S = \{\mathbf{x} : \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, where $\mathbf{A}$ is $m \times n$ and rank$(\mathbf{A}) = m < n$. $\bar{\mathbf{x}}$ is an extreme point of S if and only if $\bar{\mathbf{x}}$ is the intersection of n linearly independent hyperplanes.

only if.

Let $\bar{\mathbf{x}}$ be an extreme point of S. To get a contradiction we will assume that $\bar{\mathbf{x}}$ lies on less than n linearly independent hyperplanes. By definition of S, $\bar{\mathbf{x}}$ lies on the m linearly independent hyperplanes forming the constraint set $\mathbf{A} \cdot \bar{\mathbf{x}} = \mathbf{b}$.

Goal

First we want to develop a method of identifying the extreme points of a system of constraints in standard form.

Theorem

Let $S = \{\mathbf{x} : \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, *where* $\mathbf{A}$ *is* $m \times n$ *and* $\mathrm{rank}(\mathbf{A}) = m < n$. $\overline{\mathbf{x}}$ *is an extreme point of* S *if and only if* $\overline{\mathbf{x}}$ *is the intersection of* n *linearly independent hyperplanes.*

only if.

Let $\overline{\mathbf{x}}$ be an extreme point of S. To get a contradiction we will assume that $\overline{\mathbf{x}}$ lies on less than n linearly independent hyperplanes. By definition of S, $\overline{\mathbf{x}}$ lies on the m linearly independent hyperplanes forming the constraint set $\mathbf{A} \cdot \overline{\mathbf{x}} = \mathbf{b}$. Thus $\overline{\mathbf{x}}$ must also lie on exactly $p < n - m$ of the hyperplanes corresponding to the constraints $\mathbf{x} \geq \mathbf{0}$.

Goal

First we want to develop a method of identifying the extreme points of a system of constraints in standard form.

Theorem

Let $S = \{\mathbf{x} : \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, *where* $\mathbf{A}$ *is* $m \times n$ *and* $\text{rank}(\mathbf{A}) = m < n$. $\overline{\mathbf{x}}$ *is an extreme point of S if and only if* $\overline{\mathbf{x}}$ *is the intersection of n linearly independent hyperplanes.*

only if.

Let $\overline{\mathbf{x}}$ be an extreme point of S. To get a contradiction we will assume that $\overline{\mathbf{x}}$ lies on less than n linearly independent hyperplanes. By definition of S, $\overline{\mathbf{x}}$ lies on the m linearly independent hyperplanes forming the constraint set $\mathbf{A} \cdot \overline{\mathbf{x}} = \mathbf{b}$. Thus $\overline{\mathbf{x}}$ must also lie on exactly $p < n - m$ of the hyperplanes corresponding to the constraints $\mathbf{x} \geq \mathbf{0}$. Without loss of generality we can assume that $\overline{x}_i = 0$ for $i = 1, \ldots, p$ and $\overline{x}_i > 0$ for $i = p + 1, \ldots, n$.

Goal

First we want to develop a method of identifying the extreme points of a system of constraints in standard form.

Theorem

Let $S = \{\mathbf{x} : \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$*, where* $\mathbf{A}$ *is* $m \times n$ *and* $\text{rank}(\mathbf{A}) = m < n$. $\overline{\mathbf{x}}$ *is an extreme point of* S *if and only if* $\overline{\mathbf{x}}$ *is the intersection of* n *linearly independent hyperplanes.*

only if.

Let $\overline{\mathbf{x}}$ be an extreme point of S. To get a contradiction we will assume that $\overline{\mathbf{x}}$ lies on less than n linearly independent hyperplanes. By definition of S, $\overline{\mathbf{x}}$ lies on the m linearly independent hyperplanes forming the constraint set $\mathbf{A} \cdot \overline{\mathbf{x}} = \mathbf{b}$. Thus $\overline{\mathbf{x}}$ must also lie on exactly $p < n - m$ of the hyperplanes corresponding to the constraints $\mathbf{x} \geq \mathbf{0}$. Without loss of generality we can assume that $\overline{x}_i = 0$ for $i = 1, \ldots, p$ and $\overline{x}_i > 0$ for $i = p + 1, \ldots, n$. Thus we can create a new system of constraints $\mathbf{Qx} = \mathbf{h}$ formed by adding the constraints $x_i = 0$ for $i = 1, \ldots, p$ to $\mathbf{Ax} = \mathbf{b}$. □

only if.

As $\mathbf{Q}$ is an $(m+p) \times n$ matrix where $m+p < n$ the columns of $\mathbf{Q}$ are linearly dependent. Thus there exists $\mathbf{y} \neq \mathbf{0}$ such that $\mathbf{Q}\mathbf{y} = \mathbf{0}$.

only if.

As $\mathbf{Q}$ is an $(m+p)\times n$ matrix where $m+p<n$ the columns of $\mathbf{Q}$ are linearly dependent. Thus there exists $\mathbf{y}\neq\mathbf{0}$ such that $\mathbf{Qy}=\mathbf{0}$. Now let's consider the points $\tilde{\mathbf{x}}=\overline{\mathbf{x}}+\lambda\mathbf{y}$ and $\hat{\mathbf{x}}=\overline{\mathbf{x}}-\lambda\mathbf{y}$ where $\lambda>0$.

only if.

As $\mathbf{Q}$ is an $(m+p) \times n$ matrix where $m+p < n$ the columns of $\mathbf{Q}$ are linearly dependent. Thus there exists $\mathbf{y} \neq \mathbf{0}$ such that $\mathbf{Q}\mathbf{y} = \mathbf{0}$. Now let's consider the points $\tilde{\mathbf{x}} = \overline{\mathbf{x}} + \lambda\mathbf{y}$ and $\hat{\mathbf{x}} = \overline{\mathbf{x}} - \lambda\mathbf{y}$ where $\lambda > 0$. We have that $\mathbf{Q}\tilde{\mathbf{x}} = \mathbf{Q}(\overline{\mathbf{x}} + \lambda\mathbf{y}) = \mathbf{h} + \lambda\mathbf{0} = \mathbf{h}$ and $\mathbf{Q}\hat{\mathbf{x}} = \mathbf{Q}(\overline{\mathbf{x}} - \lambda\mathbf{y}) = \mathbf{h} - \lambda\mathbf{0} = \mathbf{h}$.

only if.

As $\mathbf{Q}$ is an $(m+p) \times n$ matrix where $m+p < n$ the columns of $\mathbf{Q}$ are linearly dependent. Thus there exists $\mathbf{y} \neq \mathbf{0}$ such that $\mathbf{Q}\mathbf{y} = \mathbf{0}$. Now let's consider the points $\tilde{\mathbf{x}} = \overline{\mathbf{x}} + \lambda\mathbf{y}$ and $\hat{\mathbf{x}} = \overline{\mathbf{x}} - \lambda\mathbf{y}$ where $\lambda > 0$. We have that $\mathbf{Q}\tilde{\mathbf{x}} = \mathbf{Q}(\overline{\mathbf{x}} + \lambda\mathbf{y}) = \mathbf{h} + \lambda\mathbf{0} = \mathbf{h}$ and $\mathbf{Q}\hat{\mathbf{x}} = \mathbf{Q}(\overline{\mathbf{x}} - \lambda\mathbf{y}) = \mathbf{h} - \lambda\mathbf{0} = \mathbf{h}$. Thus $\mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$ and $\tilde{x}_i = \hat{x}_i = 0$ for $i = 1, \ldots, p$.

only if.

As $\mathbf{Q}$ is an $(m+p) \times n$ matrix where $m+p < n$ the columns of $\mathbf{Q}$ are linearly dependent. Thus there exists $\mathbf{y} \neq \mathbf{0}$ such that $\mathbf{Qy} = \mathbf{0}$. Now let's consider the points $\tilde{\mathbf{x}} = \overline{\mathbf{x}} + \lambda \mathbf{y}$ and $\hat{\mathbf{x}} = \overline{\mathbf{x}} - \lambda \mathbf{y}$ where $\lambda > 0$. We have that $\mathbf{Q}\tilde{\mathbf{x}} = \mathbf{Q}(\overline{\mathbf{x}} + \lambda \mathbf{y}) = \mathbf{h} + \lambda \mathbf{0} = \mathbf{h}$ and $\mathbf{Q}\hat{\mathbf{x}} = \mathbf{Q}(\overline{\mathbf{x}} - \lambda \mathbf{y}) = \mathbf{h} - \lambda \mathbf{0} = \mathbf{h}$. Thus $\mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$ and $\tilde{x}_i = \hat{x}_i = 0$ for $i = 1, \ldots, p$. Since $\overline{x}_j > 0$ for $j = p+1, \ldots, n$ there exists λ such that $\tilde{x}_j = \overline{x}_j + \lambda y_j > 0$ and $\hat{x}_j = \overline{x}_j - \lambda y_j > 0$ for $j = p+1, \ldots, n$. Thus $\tilde{\mathbf{x}}, \hat{\mathbf{x}} \in S$. However $\overline{\mathbf{x}} = (1/2)\tilde{\mathbf{x}} + (1/2)\hat{\mathbf{x}}$ contradicting the fact that $\overline{\mathbf{x}}$ is an extreme point of S. Thus $\overline{\mathbf{x}}$ has to lie on n linearly independent hyperplanes. $\square$

only if.

As $\mathbf{Q}$ is an $(m+p) \times n$ matrix where $m+p<n$ the columns of $\mathbf{Q}$ are linearly dependent. Thus there exists $\mathbf{y} \neq \mathbf{0}$ such that $\mathbf{Q}\mathbf{y} = \mathbf{0}$. Now let's consider the points $\tilde{\mathbf{x}} = \overline{\mathbf{x}} + \lambda\mathbf{y}$ and $\hat{\mathbf{x}} = \overline{\mathbf{x}} - \lambda\mathbf{y}$ where $\lambda > 0$. We have that $\mathbf{Q}\tilde{\mathbf{x}} = \mathbf{Q}(\overline{\mathbf{x}} + \lambda\mathbf{y}) = \mathbf{h} + \lambda\mathbf{0} = \mathbf{h}$ and $\mathbf{Q}\hat{\mathbf{x}} = \mathbf{Q}(\overline{\mathbf{x}} - \lambda\mathbf{y}) = \mathbf{h} - \lambda\mathbf{0} = \mathbf{h}$. Thus $\mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$ and $\tilde{x}_i = \hat{x}_i = 0$ for $i = 1, \ldots, p$. Since $\overline{x}_j > 0$ for $j = p+1, \ldots, n$ there exists λ such that $\tilde{x}_j = \overline{x}_j + \lambda y_j > 0$ and $\hat{x}_j = \overline{x}_j - \lambda y_j > 0$ for $j = p+1, \ldots, n$. Thus $\tilde{\mathbf{x}}, \hat{\mathbf{x}} \in S$. However $\overline{\mathbf{x}} = (1/2)\tilde{\mathbf{x}} + (1/2)\hat{\mathbf{x}}$ contradicting the fact that $\overline{\mathbf{x}}$ is an extreme point of S. Thus $\overline{\mathbf{x}}$ has to lie on n linearly independent hyperplanes. $\square$

if.

Let $\overline{\mathbf{x}} \in S$ be the intersection of n linearly independent hyperplanes. Without loss of generality we can let these hyperplanes be denoted by $\mathbf{A}\mathbf{x} = \mathbf{b}$ and $x_i = 0$ for $i = 1, \ldots, n-m$. Now select $\tilde{\mathbf{x}}, \hat{\mathbf{x}} \in S$ and $\alpha \in (0,1)$ such that $\overline{\mathbf{x}} = \alpha\tilde{\mathbf{x}} + (1-\alpha)\hat{\mathbf{x}}$.

only if.

As $\mathbf{Q}$ is an $(m+p) \times n$ matrix where $m+p < n$ the columns of $\mathbf{Q}$ are linearly dependent. Thus there exists $\mathbf{y} \neq \mathbf{0}$ such that $\mathbf{Q}\mathbf{y} = \mathbf{0}$. Now let's consider the points $\tilde{\mathbf{x}} = \overline{\mathbf{x}} + \lambda\mathbf{y}$ and $\hat{\mathbf{x}} = \overline{\mathbf{x}} - \lambda\mathbf{y}$ where $\lambda > 0$. We have that $\mathbf{Q}\tilde{\mathbf{x}} = \mathbf{Q}(\overline{\mathbf{x}} + \lambda\mathbf{y}) = \mathbf{h} + \lambda\mathbf{0} = \mathbf{h}$ and $\mathbf{Q}\hat{\mathbf{x}} = \mathbf{Q}(\overline{\mathbf{x}} - \lambda\mathbf{y}) = \mathbf{h} - \lambda\mathbf{0} = \mathbf{h}$. Thus $\mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$ and $\tilde{x}_i = \hat{x}_i = 0$ for $i = 1, \ldots, p$. Since $\overline{x}_j > 0$ for $j = p+1, \ldots, n$ there exists λ such that $\tilde{x}_j = \overline{x}_j + \lambda y_j > 0$ and $\hat{x}_j = \overline{x}_j - \lambda y_j > 0$ for $j = p+1, \ldots, n$. Thus $\tilde{\mathbf{x}}, \hat{\mathbf{x}} \in S$. However $\overline{\mathbf{x}} = (1/2)\tilde{\mathbf{x}} + (1/2)\hat{\mathbf{x}}$ contradicting the fact that $\overline{\mathbf{x}}$ is an extreme point of S. Thus $\overline{\mathbf{x}}$ has to lie on n linearly independent hyperplanes. $\square$

if.

Let $\overline{\mathbf{x}} \in S$ be the intersection of n linearly independent hyperplanes. Without loss of generality we can let these hyperplanes be denoted by $\mathbf{A}\mathbf{x} = \mathbf{b}$ and $x_i = 0$ for $i = 1, \ldots, n-m$. Now select $\tilde{\mathbf{x}}, \hat{\mathbf{x}} \in S$ and $\alpha \in (0, 1)$ such that $\overline{\mathbf{x}} = \alpha\tilde{\mathbf{x}} + (1-\alpha)\hat{\mathbf{x}}$. We have that $\alpha\tilde{x}_i + (1-\alpha)\hat{x}_i = \overline{x}_i = 0$ for $i = 1, \ldots, n-m$. As $\alpha, (1-\alpha) > 0$ and $\tilde{x}_1, \hat{x}_i \geq 0$ we have that $\tilde{x}_i = \hat{x}_i = \overline{x}_i = 0$ for $i = 1, \ldots, n-m$. $\square$

if.

As $\overline{\mathbf{x}}, \tilde{\mathbf{x}}, \hat{\mathbf{x}} \in S$ we have that $\mathbf{A}\overline{\mathbf{x}} = \mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$.

if.

As $\overline{\mathbf{x}}, \tilde{\mathbf{x}}, \hat{\mathbf{x}} \in S$ we have that $\mathbf{A}\overline{\mathbf{x}} = \mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$. From this, and the previous result, we have that $\overline{\mathbf{x}}$, $\tilde{\mathbf{x}}$, and $\hat{\mathbf{x}}$ are each located at the intersection of the same n linearly independent hyperplanes.

if.

As $\overline{\mathbf{x}}, \tilde{\mathbf{x}}, \hat{\mathbf{x}} \in S$ we have that $\mathbf{A}\overline{\mathbf{x}} = \mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$. From this, and the previous result, we have that $\overline{\mathbf{x}}$, $\tilde{\mathbf{x}}$, and $\hat{\mathbf{x}}$ are each located at the intersection of the same n linearly independent hyperplanes. Therefore $\overline{\mathbf{x}} = \tilde{\mathbf{x}} = \hat{\mathbf{x}}$. Thus $\overline{\mathbf{x}}$ is indeed an extreme point. $\square$

if.

As $\overline{\mathbf{x}}, \tilde{\mathbf{x}}, \hat{\mathbf{x}} \in S$ we have that $\mathbf{A}\overline{\mathbf{x}} = \mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$. From this, and the previous result, we have that $\overline{\mathbf{x}}$, $\tilde{\mathbf{x}}$, and $\hat{\mathbf{x}}$ are each located at the intersection of the same n linearly independent hyperplanes. Therefore $\overline{\mathbf{x}} = \tilde{\mathbf{x}} = \hat{\mathbf{x}}$. Thus $\overline{\mathbf{x}}$ is indeed an extreme point. $\square$

Recognizing directions

Now that we have a different way of recognizing extreme points we will also determine a way of recognizing directions.

if.

As $\overline{\mathbf{x}}, \tilde{\mathbf{x}}, \hat{\mathbf{x}} \in S$ we have that $\mathbf{A}\overline{\mathbf{x}} = \mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$. From this, and the previous result, we have that $\overline{\mathbf{x}}$, $\tilde{\mathbf{x}}$, and $\hat{\mathbf{x}}$ are each located at the intersection of the same n linearly independent hyperplanes. Therefore $\overline{\mathbf{x}} = \tilde{\mathbf{x}} = \hat{\mathbf{x}}$. Thus $\overline{\mathbf{x}}$ is indeed an extreme point. $\square$

Recognizing directions

Now that we have a different way of recognizing extreme points we will also determine a way of recognizing directions.

Theorem

Let $S = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$. *Then* $\mathbf{d}$ *is a direction of* S *if and only if* $\mathbf{d} \in D = \{\mathbf{d} : \mathbf{A}\mathbf{d} = \mathbf{0}, \mathbf{d} \geq \mathbf{0}, \mathbf{d} \neq \mathbf{0}\}$.

only if.

Let $\mathbf{d}$ be a direction of S. Thus $\mathbf{d} \neq \mathbf{0}$ and $\mathbf{x} + \lambda\mathbf{d} \in S$ for all $\mathbf{x} \in S$ and $\lambda \geq 0$.

only if.

Let $\mathbf{d}$ be a direction of S. Thus $\mathbf{d} \neq \mathbf{0}$ and $\mathbf{x} + \lambda\mathbf{d} \in S$ for all $\mathbf{x} \in S$ and $\lambda \geq 0$. Thus, for $\lambda = 1$, we have that $\mathbf{b} = \mathbf{A}(\mathbf{x} + \mathbf{d}) = \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{d} = \mathbf{b} + \mathbf{A}\mathbf{d}$. Thus $\mathbf{A}\mathbf{d} = \mathbf{0}$.

only if.

Let $\mathbf{d}$ be a direction of S. Thus $\mathbf{d} \neq \mathbf{0}$ and $\mathbf{x} + \lambda\mathbf{d} \in S$ for all $\mathbf{x} \in S$ and $\lambda \geq 0$. Thus, for $\lambda = 1$, we have that $\mathbf{b} = \mathbf{A}(\mathbf{x} + \mathbf{d}) = \mathbf{Ax} + \mathbf{Ad} = \mathbf{b} + \mathbf{Ad}$. Thus $\mathbf{Ad} = \mathbf{0}$. We also have that $\mathbf{x} + \lambda\mathbf{d} \geq 0$. This implies that $\mathbf{d} \geq \mathbf{0}$ because otherwise λ could be made large enough to make $\mathbf{x} + \lambda\mathbf{d} < \mathbf{0}$.

only if.

Let $\mathbf{d}$ be a direction of S. Thus $\mathbf{d} \neq \mathbf{0}$ and $\mathbf{x} + \lambda\mathbf{d} \in S$ for all $\mathbf{x} \in S$ and $\lambda \geq 0$. Thus, for $\lambda = 1$, we have that $\mathbf{b} = \mathbf{A}(\mathbf{x} + \mathbf{d}) = \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{d} = \mathbf{b} + \mathbf{A}\mathbf{d}$. Thus $\mathbf{A}\mathbf{d} = \mathbf{0}$. We also have that $\mathbf{x} + \lambda\mathbf{d} \geq 0$. This implies that $\mathbf{d} \geq \mathbf{0}$ because otherwise λ could be made large enough to make $\mathbf{x} + \lambda\mathbf{d} < \mathbf{0}$. Thus $\mathbf{d} \in D$. $\square$

only if.

Let $\mathbf{d}$ be a direction of S. Thus $\mathbf{d} \neq \mathbf{0}$ and $\mathbf{x} + \lambda\mathbf{d} \in S$ for all $\mathbf{x} \in S$ and $\lambda \geq 0$. Thus, for $\lambda = 1$, we have that $\mathbf{b} = \mathbf{A}(\mathbf{x} + \mathbf{d}) = \mathbf{Ax} + \mathbf{Ad} = \mathbf{b} + \mathbf{Ad}$. Thus $\mathbf{Ad} = \mathbf{0}$. We also have that $\mathbf{x} + \lambda\mathbf{d} \geq 0$. This implies that $\mathbf{d} \geq \mathbf{0}$ because otherwise λ could be made large enough to make $\mathbf{x} + \lambda\mathbf{d} < \mathbf{0}$. Thus $\mathbf{d} \in D$. □

if.

Let $\mathbf{d} \in D$ and let $\mathbf{x} \in S$.

only if.

Let $\mathbf{d}$ be a direction of S. Thus $\mathbf{d} \neq \mathbf{0}$ and $\mathbf{x} + \lambda\mathbf{d} \in S$ for all $\mathbf{x} \in S$ and $\lambda \geq 0$. Thus, for $\lambda = 1$, we have that $\mathbf{b} = \mathbf{A}(\mathbf{x} + \mathbf{d}) = \mathbf{Ax} + \mathbf{Ad} = \mathbf{b} + \mathbf{Ad}$. Thus $\mathbf{Ad} = \mathbf{0}$. We also have that $\mathbf{x} + \lambda\mathbf{d} \geq 0$. This implies that $\mathbf{d} \geq \mathbf{0}$ because otherwise λ could be made large enough to make $\mathbf{x} + \lambda\mathbf{d} < \mathbf{0}$. Thus $\mathbf{d} \in D$. □

if.

Let $\mathbf{d} \in D$ and let $\mathbf{x} \in S$. We have that, by definition of D, $\mathbf{d} \neq \mathbf{0}$.

only if.

Let $\mathbf{d}$ be a direction of S. Thus $\mathbf{d} \neq \mathbf{0}$ and $\mathbf{x} + \lambda\mathbf{d} \in S$ for all $\mathbf{x} \in S$ and $\lambda \geq 0$. Thus, for $\lambda = 1$, we have that $\mathbf{b} = \mathbf{A}(\mathbf{x} + \mathbf{d}) = \mathbf{Ax} + \mathbf{Ad} = \mathbf{b} + \mathbf{Ad}$. Thus $\mathbf{Ad} = \mathbf{0}$. We also have that $\mathbf{x} + \lambda\mathbf{d} \geq 0$. This implies that $\mathbf{d} \geq \mathbf{0}$ because otherwise λ could be made large enough to make $\mathbf{x} + \lambda\mathbf{d} < \mathbf{0}$. Thus $\mathbf{d} \in D$. $\square$

if.

Let $\mathbf{d} \in D$ and let $\mathbf{x} \in S$. We have that, by definition of D, $\mathbf{d} \neq \mathbf{0}$. Also by the definition of D we have that $\mathbf{Ad} = \mathbf{0}$, thus for $\lambda \geq 0$ we have that $\mathbf{A}(\mathbf{x} + \lambda\mathbf{d}) = \mathbf{Ax} + \mathbf{Ad} = \mathbf{b}$ also as $\mathbf{x}, \mathbf{d} \geq \mathbf{0}$ we have that $\mathbf{x} + \lambda\mathbf{d} \geq \mathbf{0}$.

only if.

Let $\mathbf{d}$ be a direction of S. Thus $\mathbf{d} \neq \mathbf{0}$ and $\mathbf{x} + \lambda\mathbf{d} \in S$ for all $\mathbf{x} \in S$ and $\lambda \geq 0$. Thus, for $\lambda = 1$, we have that $\mathbf{b} = \mathbf{A}(\mathbf{x} + \mathbf{d}) = \mathbf{Ax} + \mathbf{Ad} = \mathbf{b} + \mathbf{Ad}$. Thus $\mathbf{Ad} = \mathbf{0}$. We also have that $\mathbf{x} + \lambda\mathbf{d} \geq 0$. This implies that $\mathbf{d} \geq \mathbf{0}$ because otherwise λ could be made large enough to make $\mathbf{x} + \lambda\mathbf{d} < \mathbf{0}$. Thus $\mathbf{d} \in D$. □

if.

Let $\mathbf{d} \in D$ and let $\mathbf{x} \in S$. We have that, by definition of D, $\mathbf{d} \neq \mathbf{0}$. Also by the definition of D we have that $\mathbf{Ad} = \mathbf{0}$, thus for $\lambda \geq 0$ we have that $\mathbf{A}(\mathbf{x} + \lambda\mathbf{d}) = \mathbf{Ax} + \mathbf{Ad} = \mathbf{b}$ also as $\mathbf{x}, \mathbf{d} \geq \mathbf{0}$ we have that $\mathbf{x} + \lambda\mathbf{d} \geq \mathbf{0}$. Thus for all $\lambda \geq 0$ $\mathbf{x} + \lambda\mathbf{d} \in S$. Thus $\mathbf{d}$ is a direction of S □

Bounded and unbounded sets

Before continuing on, we define what it it means for a set, specifically a subset of $\mathbb{R}^n$ to be either bounded or unbounded.

Bounded and unbounded sets

Before continuing on, we define what it it means for a set, specifically a subset of $\mathbb{R}^n$ to be either bounded or unbounded.

Definition (Bounded Set)

A subset S of $\mathbb{R}^n$ is bounded if it can be contained within an n-dimensional ball. Mathematically S is bounded if there exist a point $\bar{\mathbf{x}} \in S$ and a radius $r > 0$ such that for any point $\mathbf{x} \in S$ the distance from $\bar{\mathbf{x}}$ to $\mathbf{x}$ is less than r.

Bounded and unbounded sets

Before continuing on, we define what it it means for a set, specifically a subset of $\mathbb{R}^n$ to be either bounded or unbounded.

Definition (Bounded Set)

A subset S of $\mathbb{R}^n$ is bounded if it can be contained within an n-dimensional ball. Mathematically S is bounded if there exist a point $\bar{\mathbf{x}} \in S$ and a radius $r > 0$ such that for any point $\mathbf{x} \in S$ the distance from $\bar{\mathbf{x}}$ to $\mathbf{x}$ is less than r.

Definition (Unbounded Set)

An unbounded set is simply a set which is not bounded.

Outline

Importance

Now we will see the importance of considering extreme points and extreme directions when considering a linear program. However first we will see how a set S relates to its extreme points and extreme directions.

Importance

Now we will see the importance of considering extreme points and extreme directions when considering a linear program. However first we will see how a set S relates to its extreme points and extreme directions.

Theorem

Let $S = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ be non-empty, and let E be the set of extreme points of S and let D be the set of all extreme directions of S. Then:

Importance

Now we will see the importance of considering extreme points and extreme directions when considering a linear program. However first we will see how a set S relates to its extreme points and extreme directions.

Theorem

Let $S = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ *be non-empty, and let E be the set of extreme points of S and let D be the set of all extreme directions of S. Then:*

1. *S has at least one extreme point and at most a finite number of extreme points, thus* $E = \{\mathbf{x}_1, \ldots, \mathbf{x}_p\} \neq \emptyset$.

Importance

Now we will see the importance of considering extreme points and extreme directions when considering a linear program. However first we will see how a set S relates to its extreme points and extreme directions.

Theorem

Let $S = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ be non-empty, and let E be the set of extreme points of S and let D be the set of all extreme directions of S. Then:

1. *S has at least one extreme point and at most a finite number of extreme points, thus $E = \{\mathbf{x}_1, \ldots, \mathbf{x}_p\} \neq \emptyset$.*
2. *S is unbounded of and only if S has at least one extreme direction.*

Importance

Now we will see the importance of considering extreme points and extreme directions when considering a linear program. However first we will see how a set S relates to its extreme points and extreme directions.

Theorem

Let $S = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ be non-empty, and let E be the set of extreme points of S and let D be the set of all extreme directions of S. Then:

1. *S has at least one extreme point and at most a finite number of extreme points, thus $E = \{\mathbf{x}_1, \ldots, \mathbf{x}_p\} \neq \emptyset$.*
2. *S is unbounded of and only if S has at least one extreme direction.*
3. *if S is unbounded then S has a finite number of extreme directions, $D = \{\mathbf{d}_1, \ldots, \mathbf{d}_q\} \neq \emptyset$.*

Importance

Now we will see the importance of considering extreme points and extreme directions when considering a linear program. However first we will see how a set S relates to its extreme points and extreme directions.

Theorem

Let $S = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ be non-empty, and let E be the set of extreme points of S and let D be the set of all extreme directions of S. Then:

1. *S has at least one extreme point and at most a finite number of extreme points, thus $E = \{\mathbf{x}_1, \ldots, \mathbf{x}_p\} \neq \emptyset$.*
2. *S is unbounded of and only if S has at least one extreme direction.*
3. *if S is unbounded then S has a finite number of extreme directions, $D = \{\mathbf{d}_1, \ldots, \mathbf{d}_q\} \neq \emptyset$.*
4. *if $\mathbf{x} \in S$, then $\mathbf{x}$ can be written as a convex combination of extreme points plus a positive combination of extreme vectors, that is $\mathbf{x} = \sum_{i=1}^{p}(\alpha_i \mathbf{x}_i) + \sum_{j=1}^{q}(\lambda_j \mathbf{d}_j)$ for $\sum_{i=1}^{p} \alpha_i = 1$, $\alpha_i \geq 0$, $\lambda_j \geq 0$.*

Theorem

Let $S = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ *and consider the following linear program.*

$$\begin{array}{c} \text{maximize } z = \mathbf{c}\mathbf{x} \\ \text{subject to } \mathbf{x} \in S. \end{array}$$

Suppose S *is unbounded and has extreme points* $E = \{\mathbf{x}_1, \ldots, \mathbf{x}_p\} \neq \emptyset$ *and extreme directions* $D = \{\mathbf{d}_1, \ldots, \mathbf{d}_q\} \neq \emptyset$*. Let* z^* *represent the optimal objective value of the linear program. Then* z^* *is finite if and only if* $\mathbf{c}\mathbf{d}_j \leq 0$ *for all* $\mathbf{d}_j \in D$*. And if a finite optimal solution exists then an extreme point optimal solution exists.*

Theorem

Let $S = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ and consider the following linear program.

$$\begin{array}{c} \textit{maximize } z = \mathbf{c}\mathbf{x} \\ \textit{subject to } \mathbf{x} \in S. \end{array}$$

Suppose S is unbounded and has extreme points $E = \{\mathbf{x}_1, \ldots, \mathbf{x}_p\} \neq \emptyset$ and extreme directions $D = \{\mathbf{d}_1, \ldots, \mathbf{d}_q\} \neq \emptyset$. Let z^* represent the optimal objective value of the linear program. Then z^* is finite if and only if $\mathbf{c}\mathbf{d}_j \leq 0$ for all $\mathbf{d}_j \in D$. And if a finite optimal solution exists then an extreme point optimal solution exists.

Proof.

By the fourth part of the previous theorem we can express any $\mathbf{x} \in S$ in terms of extreme points and extreme directions. Thus we can modify the linear program to

$$\begin{array}{c} \text{maximize } z = \sum_{i=1}^{p}(\alpha_i \mathbf{c}\mathbf{x}_i) + \sum_{j=1}^{q}(\lambda_j \mathbf{c}\mathbf{d}_j) \\ \text{subject to } \sum_{i=1}^{p} \alpha_i = 1,\ \alpha_i \geq 0,\ \lambda_j \geq 0. \end{array}$$

□

only if.

Suppose, to obtain a contradiction, that for some $\mathbf{d}_j \in D$ $\mathbf{cd} > 0$. However $\lambda_j \mathbf{cd}$ is a term of z thus z can be made arbitrarily large by making λ_j arbitrarily large.

only if.

Suppose, to obtain a contradiction, that for some $\mathbf{d}_j \in D$ $\mathbf{cd} > 0$. However $\lambda_j \mathbf{cd}$ is a term of z thus z can be made arbitrarily large by making λ_j arbitrarily large. This contradicts that z^* is finite and so $\mathbf{cd} \leq 0$ for all $\mathbf{d}_j \in D$. $\square$

only if.

Suppose, to obtain a contradiction, that for some $\mathbf{d}_j \in D$ $\mathbf{cd} > 0$. However $\lambda_j \mathbf{cd}$ is a term of z thus z can be made arbitrarily large by making λ_j arbitrarily large. This contradicts that z^* is finite and so $\mathbf{cd} \leq 0$ for all $\mathbf{d}_j \in D$. □

if.

If $\mathbf{cd} \leq 0$ for all $\mathbf{d}_j \in D$ then, since we are maximizing z, the optimum solution occurs when $\lambda_j = 0$ for all $j = 1, \ldots, q$.

only if.

Suppose, to obtain a contradiction, that for some $\mathbf{d}_j \in D$ $\mathbf{cd} > 0$. However $\lambda_j \mathbf{cd}$ is a term of z thus z can be made arbitrarily large by making λ_j arbitrarily large. This contradicts that z^* is finite and so $\mathbf{cd} \leq 0$ for all $\mathbf{d}_j \in D$. □

if.

If $\mathbf{cd} \leq 0$ for all $\mathbf{d}_j \in D$ then, since we are maximizing z, the optimum solution occurs when $\lambda_j = 0$ for all $j = 1, \ldots, q$. Thus the linear program can be simplified to can be simplified to

$$\text{maximize } z = \sum_{i=1}^{p} (\alpha_i \mathbf{c}\mathbf{x}_i)$$
$$\text{subject to } \sum_{i=1}^{p} \alpha_i = 1,\ \alpha_i \geq 0.$$

only if.

Suppose, to obtain a contradiction, that for some $\mathbf{d}_j \in D$ $\mathbf{cd} > 0$. However $\lambda_j \mathbf{cd}$ is a term of z thus z can be made arbitrarily large by making λ_j arbitrarily large. This contradicts that z^* is finite and so $\mathbf{cd} \leq 0$ for all $\mathbf{d}_j \in D$. □

if.

If $\mathbf{cd} \leq 0$ for all $\mathbf{d}_j \in D$ then, since we are maximizing z, the optimum solution occurs when $\lambda_j = 0$ for all $j = 1, \ldots, q$. Thus the linear program can be simplified to can be simplified to

$$\text{maximize } z = \sum_{i=1}^{p} (\alpha_i \mathbf{cx}_i)$$
$$\text{subject to } \sum_{i=1}^{p} \alpha_i = 1,\ \alpha_i \geq 0.$$

As there are finitely many extreme points there must be an extreme point $\mathbf{x}_k$ for which $(cx)_k \geq (cx)_i$ for all $i = 1, \ldots, p$.

only if.

Suppose, to obtain a contradiction, that for some $\mathbf{d}_j \in D$ $\mathbf{cd} > 0$. However $\lambda_j \mathbf{cd}$ is a term of z thus z can be made arbitrarily large by making λ_j arbitrarily large. This contradicts that z^* is finite and so $\mathbf{cd} \leq 0$ for all $\mathbf{d}_j \in D$. □

if.

If $\mathbf{cd} \leq 0$ for all $\mathbf{d}_j \in D$ then, since we are maximizing z, the optimum solution occurs when $\lambda_j = 0$ for all $j = 1, \ldots, q$. Thus the linear program can be simplified to can be simplified to

$$\text{maximize } z = \sum_{i=1}^{p} (\alpha_i \mathbf{cx}_i)$$
$$\text{subject to } \sum_{i=1}^{p} \alpha_i = 1,\ \alpha_i \geq 0.$$

As there are finitely many extreme points there must be an extreme point $\mathbf{x}_k$ for which $(cx)_k \geq (cx)_i$ for all $i = 1, \ldots, p$. Thus we have that $z = \sum_{i=1}^{p} (\alpha_i \mathbf{cx}_i) \leq \sum_{i=1}^{p} (\alpha_i \mathbf{cx}_k) = \mathbf{cx}_k$. Thus z^* is finite and in fact $z^* = \mathbf{cx}_k$ where $\mathbf{x}_k$ is the optimal extreme point. □

Outline

goal

We have now shown that we can solve linear programs by restricting ourselves to the extreme points of the feasible space. However we still need to develop a way of finding these extreme point non-graphically.

goal

We have now shown that we can solve linear programs by restricting ourselves to the extreme points of the feasible space. However we still need to develop a way of finding these extreme point non-graphically.

Finding basic feasible solutions

Consider a linear system of equations $\mathbf{Ax} = \mathbf{b}$. Where $\mathbf{A}$ is an $m \times n$ matrix $\mathbf{b} = (b_1, \ldots, b_m)^t$, and $\mathbf{x} = (x_1, \ldots, x_n)^t$.

goal

We have now shown that we can solve linear programs by restricting ourselves to the extreme points of the feasible space. However we still need to develop a way of finding these extreme point non-graphically.

Finding basic feasible solutions

Consider a linear system of equations $\mathbf{A}\mathbf{x} = \mathbf{b}$. Where $\mathbf{A}$ is an $m \times n$ matrix $\mathbf{b} = (b_1, \ldots, b_m)^t$, and $\mathbf{x} = (x_1, \ldots, x_n)^t$. We will assume that $rank(\mathbf{A}) = m \leq n$. That is we assume that the rows of $\mathbf{A}$ are linearly independent.

goal

We have now shown that we can solve linear programs by restricting ourselves to the extreme points of the feasible space. However we still need to develop a way of finding these extreme point non-graphically.

Finding basic feasible solutions

Consider a linear system of equations $\mathbf{Ax} = \mathbf{b}$. Where $\mathbf{A}$ is an $m \times n$ matrix $\mathbf{b} = (b_1, \ldots, b_m)^t$, and $\mathbf{x} = (x_1, \ldots, x_n)^t$. We will assume that $rank(\mathbf{A}) = m \leq n$. That is we assume that the rows of $\mathbf{A}$ are linearly independent. We also assume that the columns of $\mathbf{A}$ can be rearranged so that $\mathbf{A}$ can be written as $\mathbf{A} = (\mathbf{B} : \mathbf{N})$. Where $\mathbf{B}$ is a nonsingular $m \times m$ matrix. We will refer to $\mathbf{B}$ as the *basis matrix*.

Finding basic feasible solutions

Thus we can rewrite $\mathbf{Ax} = \mathbf{b}$ as $\mathbf{Bx}_B + \mathbf{Nx}_N = \mathbf{b}$ where $\mathbf{x} = \binom{\mathbf{x}_B}{\mathbf{x}_N}$.

Finding basic feasible solutions

Thus we can rewrite $\mathbf{Ax} = \mathbf{b}$ as $\mathbf{Bx}_B + \mathbf{Nx}_N = \mathbf{b}$ where $\mathbf{x} = \binom{\mathbf{x}_B}{\mathbf{x}_N}$. As $\mathbf{B}$ is non-singular the inverse of $\mathbf{B}$ exists. Thus we have that $\mathbf{B}^{-1}\mathbf{Bx}_B + \mathbf{B}^{-1}\mathbf{Nx}_N = \mathbf{B}^{-1}\mathbf{b}$. Which is equivalent to sating that $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{Nx}_N$.

Finding basic feasible solutions

Thus we can rewrite $\mathbf{Ax} = \mathbf{b}$ as $\mathbf{B}\mathbf{x}_B + \mathbf{N}\mathbf{x}_N = \mathbf{b}$ where $\mathbf{x} = \binom{\mathbf{x}_B}{\mathbf{x}_N}$. As $\mathbf{B}$ is non-singular the inverse of $\mathbf{B}$ exists. Thus we have that $\mathbf{B}^{-1}\mathbf{B}\mathbf{x}_B + \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_N = \mathbf{B}^{-1}\mathbf{b}$. Which is equivalent to sating that $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_N$. If we set $\mathbf{x}_N = \mathbf{0}$ then $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b}$ and $\mathbf{x} = \binom{\mathbf{B}^{-1}\mathbf{b}}{\mathbf{0}}$. This value is called a *basic solution*. We refer to the $\mathbf{x}_B$ as the vector of *basic variables* and we refer to $\mathbf{x}_N$ as the vector of *nonbasic variables*.

Finding basic feasible solutions

Thus we can rewrite $\mathbf{Ax} = \mathbf{b}$ as $\mathbf{Bx}_B + \mathbf{Nx}_N = \mathbf{b}$ where $\mathbf{x} = \binom{\mathbf{x}_B}{\mathbf{x}_N}$. As $\mathbf{B}$ is non-singular the inverse of $\mathbf{B}$ exists. Thus we have that $\mathbf{B}^{-1}\mathbf{Bx}_B + \mathbf{B}^{-1}\mathbf{Nx}_N = \mathbf{B}^{-1}\mathbf{b}$. Which is equivalent to sating that $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{Nx}_N$. If we set $\mathbf{x}_N = \mathbf{0}$ then $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b}$ and $\mathbf{x} = \binom{\mathbf{B}^{-1}\mathbf{b}}{\mathbf{0}}$. This value is called a *basic solution*. We refer to the $\mathbf{x}_B$ as the vector of *basic variables* and we refer to $\mathbf{x}_N$ as the vector of *nonbasic variables*. If $\mathbf{x} \geq \mathbf{0}$ then $\mathbf{x}$ is called a *basic feasible solution*.

Finding basic feasible solutions

Thus we can rewrite $\mathbf{Ax} = \mathbf{b}$ as $\mathbf{Bx}_B + \mathbf{Nx}_N = \mathbf{b}$ where $\mathbf{x} = \binom{\mathbf{x}_B}{\mathbf{x}_N}$. As $\mathbf{B}$ is non-singular the inverse of $\mathbf{B}$ exists. Thus we have that $\mathbf{B}^{-1}\mathbf{Bx}_B + \mathbf{B}^{-1}\mathbf{Nx}_N = \mathbf{B}^{-1}\mathbf{b}$. Which is equivalent to sating that $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{Nx}_N$. If we set $\mathbf{x}_N = \mathbf{0}$ then $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b}$ and $\mathbf{x} = \binom{\mathbf{B}^{-1}\mathbf{b}}{\mathbf{0}}$. This value is called a *basic solution*. We refer to the $\mathbf{x}_B$ as the vector of *basic variables* and we refer to $\mathbf{x}_N$ as the vector of *nonbasic variables*. If $\mathbf{x} \geq \mathbf{0}$ then $\mathbf{x}$ is called a *basic feasible solution*. Similarly if $\mathbf{x} = \mathbf{0}$ then the basic solution is said to be *degenerate* otherwise it is *non-degenerate*.

Outline

Theorem

Let $S = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$*, where* $\mathbf{A}$ *is* $m \times n$ *and* $\text{rank}(\mathbf{A}) = m < n$*.* $\mathbf{x}$ *is an extreme point of* S *if and only if* $\mathbf{x}$ *is a basic feasible solution.*

Theorem

Let $S = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, where $\mathbf{A}$ is $m \times n$ and $rank(\mathbf{A}) = m < n$. $\mathbf{x}$ is an extreme point of S if and only if $\mathbf{x}$ is a basic feasible solution.

only if.

Let $\mathbf{x} \in S$ be an extreme point.

Theorem

Let $S = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, where $\mathbf{A}$ is $m \times n$ and $rank(\mathbf{A}) = m < n$. $\mathbf{x}$ is an extreme point of S if and only if $\mathbf{x}$ is a basic feasible solution.

only if.

Let $\mathbf{x} \in S$ be an extreme point. Thus $\mathbf{x}$ is the intersection of n linearly independent hyperplanes. From the definition of S the constraints $\mathbf{A}\mathbf{x} = \mathbf{b}$ provide m of these hyperplanes. Thus the remaining $n - m$ hyperplanes must come from the non-negativity constraints.

Theorem

Let $S = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, where $\mathbf{A}$ is $m \times n$ and $rank(\mathbf{A}) = m < n$. $\mathbf{x}$ is an extreme point of S if and only if $\mathbf{x}$ is a basic feasible solution.

only if.

Let $\mathbf{x} \in S$ be an extreme point. Thus $\mathbf{x}$ is the intersection of n linearly independent hyperplanes. From the definition of S the constraints $\mathbf{A}\mathbf{x} = \mathbf{b}$ provide m of these hyperplanes. Thus the remaining $n - m$ hyperplanes must come from the non-negativity constraints. Thus at least $n - m$ of these constraints are satisfied as equalities by the extreme point $\mathbf{x}$. Thus $\mathbf{x}_N = \mathbf{0}$ can be used to represent $m - n$ such constraints.

Theorem

Let $S = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, where $\mathbf{A}$ is $m \times n$ and $rank(\mathbf{A}) = m < n$. $\mathbf{x}$ is an extreme point of S if and only if $\mathbf{x}$ is a basic feasible solution.

only if.

Let $\mathbf{x} \in S$ be an extreme point. Thus $\mathbf{x}$ is the intersection of n linearly independent hyperplanes. From the definition of S the constraints $\mathbf{A}\mathbf{x} = \mathbf{b}$ provide m of these hyperplanes. Thus the remaining $n - m$ hyperplanes must come from the non-negativity constraints. Thus at least $n - m$ of these constraints are satisfied as equalities by the extreme point $\mathbf{x}$. Thus $\mathbf{x}_N = \mathbf{0}$ can be used to represent $m - n$ such constraints. Thus $\mathbf{x}$ is the unique solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$, $\mathbf{x}_N = \mathbf{0}$. Now $\mathbf{x}_B$ is used to denote the remaining components of $\mathbf{x}$ and the matrix $\mathbf{A} = (\mathbf{B} : \mathbf{N})$ is partitioned appropriately.

Theorem

Let $S = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, where $\mathbf{A}$ is $m \times n$ and $rank(\mathbf{A}) = m < n$. $\mathbf{x}$ is an extreme point of S if and only if $\mathbf{x}$ is a basic feasible solution.

only if.

Let $\mathbf{x} \in S$ be an extreme point. Thus $\mathbf{x}$ is the intersection of n linearly independent hyperplanes. From the definition of S the constraints $\mathbf{A}\mathbf{x} = \mathbf{b}$ provide m of these hyperplanes. Thus the remaining $n - m$ hyperplanes must come from the non-negativity constraints. Thus at least $n - m$ of these constraints are satisfied as equalities by the extreme point $\mathbf{x}$. Thus $\mathbf{x}_N = \mathbf{0}$ can be used to represent $m - n$ such constraints. Thus $\mathbf{x}$ is the unique solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$, $\mathbf{x}_N = \mathbf{0}$. Now $\mathbf{x}_B$ is used to denote the remaining components of $\mathbf{x}$ and the matrix $\mathbf{A} = (\mathbf{B} : \mathbf{N})$ is partitioned appropriately. Thus $\mathbf{x}$ is the unique solution to $\mathbf{B}\mathbf{x}_B + \mathbf{N}\mathbf{x}_N = \mathbf{b}$, $\mathbf{x}_N = \mathbf{0}$. □

only if.

Thus $\mathbf{x}_B$ is the unique solution to $\mathbf{B}\mathbf{x}_B = \mathbf{b}$. Thus $\mathbf{B}$ is invertible and a basis matrix. Thus $\mathbf{x} \geq 0$ is a basic feasible solution. □

only if.

Thus $\mathbf{x}_B$ is the unique solution to $\mathbf{B}\mathbf{x}_B = \mathbf{b}$. Thus $\mathbf{B}$ is invertible and a basis matrix. Thus $\mathbf{x} \geq 0$ is a basic feasible solution. $\square$

if.

Let $\mathbf{x}$ be a basic feasible solution. Thus there exists a basis matrix $\mathbf{B}$ such that $\mathbf{x} = \binom{\mathbf{x}_B}{\mathbf{x}_N} = \binom{\mathbf{B}^{-1}\mathbf{b}}{\mathbf{0}}$.

only if.

Thus $\mathbf{x}_B$ is the unique solution to $\mathbf{B}\mathbf{x}_B = \mathbf{b}$. Thus $\mathbf{B}$ is invertible and a basis matrix. Thus $\mathbf{x} \geq 0$ is a basic feasible solution. $\square$

if.

Let $\mathbf{x}$ be a basic feasible solution. Thus there exists a basis matrix $\mathbf{B}$ such that $\mathbf{x} = \binom{\mathbf{x}_B}{\mathbf{x}_N} = \binom{\mathbf{B}^{-1}\mathbf{b}}{\mathbf{0}}$. Thus $\mathbf{x}$ is the unique solution to the system $\mathbf{A}\mathbf{x} = \mathbf{B}\mathbf{x}_B + \mathbf{N}\mathbf{x}_N = \mathbf{b}$, $\mathbf{x}_N = \mathbf{0}$. We have that $\mathbf{A}\mathbf{x} = \mathbf{b}$ represents m hyperplanes and $\mathbf{x}_N = \mathbf{0}$ represents the remaining $n - m$ hyperplanes.

only if.

Thus $\mathbf{x}_B$ is the unique solution to $\mathbf{B}\mathbf{x}_B = \mathbf{b}$. Thus $\mathbf{B}$ is invertible and a basis matrix. Thus $\mathbf{x} \geq 0$ is a basic feasible solution. $\square$

if.

Let $\mathbf{x}$ be a basic feasible solution. Thus there exists a basis matrix $\mathbf{B}$ such that $\mathbf{x} = \binom{\mathbf{x}_B}{\mathbf{x}_N} = \binom{\mathbf{B}^{-1}\mathbf{b}}{\mathbf{0}}$. Thus $\mathbf{x}$ is the unique solution to the system $\mathbf{A}\mathbf{x} = \mathbf{B}\mathbf{x}_B + \mathbf{N}\mathbf{x}_N = \mathbf{b}$, $\mathbf{x}_N = \mathbf{0}$. We have that $\mathbf{A}\mathbf{x} = \mathbf{b}$ represents m hyperplanes and $\mathbf{x}_N = \mathbf{0}$ represents the remaining $n - m$ hyperplanes. Thus $\mathbf{x}$ is the intersection of n linearly independent hyperplanes and is an extreme point. $\square$

only if.

Thus $\mathbf{x}_B$ is the unique solution to $\mathbf{B}\mathbf{x}_B = \mathbf{b}$. Thus $\mathbf{B}$ is invertible and a basis matrix. Thus $\mathbf{x} \geq 0$ is a basic feasible solution. $\square$

if.

Let $\mathbf{x}$ be a basic feasible solution. Thus there exists a basis matrix $\mathbf{B}$ such that $\mathbf{x} = \binom{\mathbf{x}_B}{\mathbf{x}_N} = \binom{\mathbf{B}^{-1}\mathbf{b}}{\mathbf{0}}$. Thus $\mathbf{x}$ is the unique solution to the system $\mathbf{A}\mathbf{x} = \mathbf{B}\mathbf{x}_B + \mathbf{N}\mathbf{x}_N = \mathbf{b}$, $\mathbf{x}_N = \mathbf{0}$. We have that $\mathbf{A}\mathbf{x} = \mathbf{b}$ represents m hyperplanes and $\mathbf{x}_N = \mathbf{0}$ represents the remaining $n - m$ hyperplanes. Thus $\mathbf{x}$ is the intersection of n linearly independent hyperplanes and is an extreme point. $\square$

All together

All together this gives us the basics of a method for finding the optimal solutions to a linear program.