COMP7180: Quantitative Methods for DAAI

(Credits from Prof. Andrew Ng)

(Credits from HKBU)

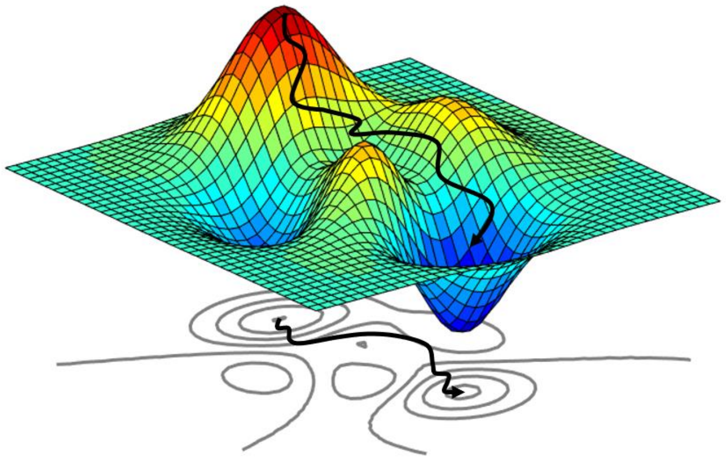
Course Instructors: Dr. Yang Liu and Dr. Bo Han

Teaching Assistant: Mr. Minghao Li

Course Contents

- Continuous and Discrete Random Variables (Week 7)
- Conditional Probability and Independence (Week 8)
- Maximum Likelihood Estimation (Week 9)
- $\boxed{\text{Mathematical Optimization (Week 10)}}$ Our Focus
- Convex and Non-Convex Optimization (Week 11)
- Quiz and Course Review (Week 12)

Mathematical Optimization



- What is mathematical optimization?

1. "Optimization" comes from the same root as "Optimal", which means best. When you optimize something, you are "making it best".

2. But "best" can vary. If you're a football player, you might want to maximize your running yards (跑场), and also minimize your fumbles (掉球率). Both maximizing and minimizing are types of optimization problems.

Credits from Dr. Julia Roberts and Dr. Mykel Kochenderfer

Mathematical Optimization

- Why do we care about Optimization ?

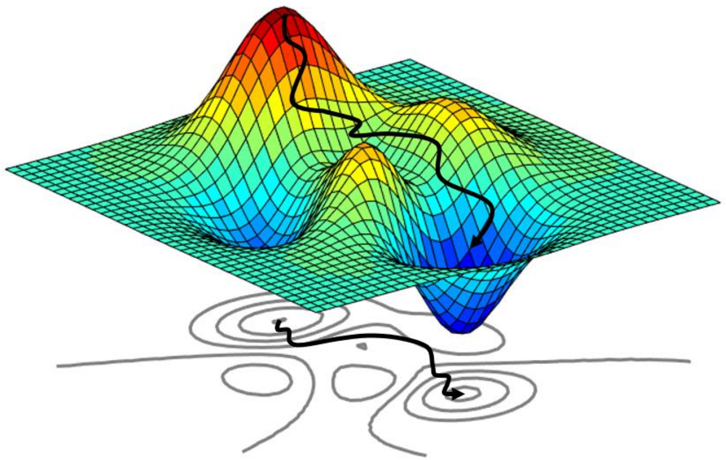
Optimization is at the heart of many AI and machine learning algorithms!

- Maximum likelihood estimation:

  Maximize $\sum_{i=1}^{n} \log P(X = x_i; \alpha)$

- Linear regression:

  Minimize $||\mathbf{X}\alpha - \mathbf{Y}||$



Credits from Dr. Yang Liu

Mathematical Optimization: Example

Let $f(x) = x^4 + 7x^3 + 5x^2 - 17x + 3$.

Then the Optimization problem is Minimize $f(x)$

How to solve it?

- Do you remember how to address the Maximum likelihood estimation? We can use the same method.

Mathematical Optimization: Example

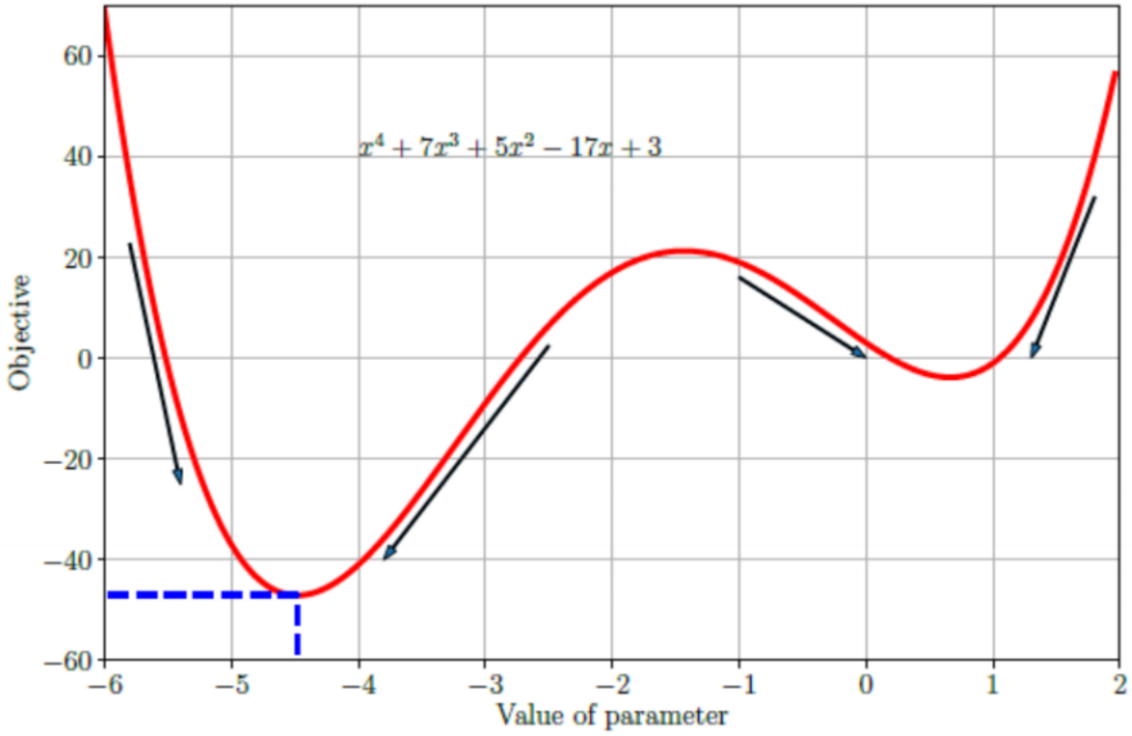
Compute the Differentiation of f:

$$\frac{df}{dx} = 4x^3 + 21x^2 + 10x - 17 = 0.$$

Then x = -4.5, -1.4, 0.7

Because x=-4.5, we have the mini value.

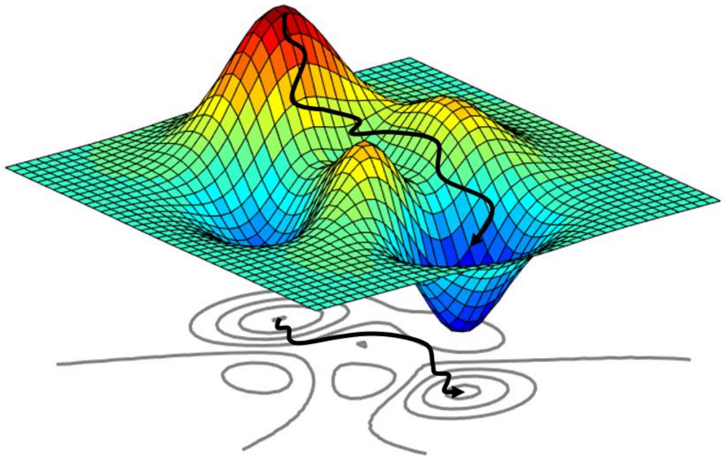
So the solution is x=-4.5.



Credits from Dr. Yang Liu

Mathematical Optimization in the “Real World”

- Mathematical Optimization is a branch of applied mathematics which is useful in many different fields. Here are a few examples:



Manufacturing

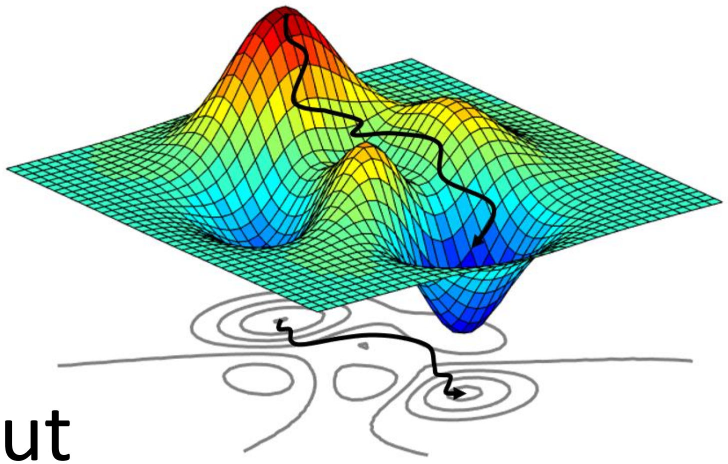
- Production
- Inventory control
- Transportation
- Scheduling
- Networks
- Finance

Engineering

- Mechanics
- Economics
- Control engineering
- Marketing
- Policy Modeling

Credits from Dr. Julia Roberts and Dr. Mykel Kochenderfer

Optimization Vocabulary



- Your basic optimization problem consists of…

1. The objective function, f(x), which is the output you're trying to maximize or minimize.
2. Variables, x_1, x_2, x_3 and so on, which are the inputs things you can control. They are abbreviated x_n to refer to individuals or x to refer to them as a group.
3. Constraints, which are equations that place limits on how big or small some variables can get. Equality constraints are usually noted $h_n(x)$ and inequality constraints are noted $g_n(x)$.

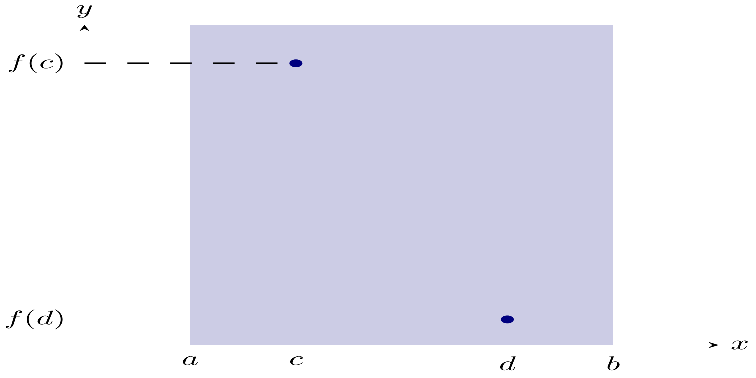
Credits from Dr. Julia Roberts and Dr. Mykel Kochenderfer

Optimization Vocabulary

- A football coach is planning practices for his running backs (跑卫).

1. His main goal is to maximize running yards (跑场) – this will become his objective function.
2. He can make his athletes spend practice time in the weight room (举重房); running sprints (跑步冲刺); or practicing ball protection (带球防卫). The amount of time spent on each is a variable.
3. However, there are limits to the total amount of time he has. Also, if he completely sacrifices ball protection he may see running yards go up, but also fumbles, so he may place an upper limit on the amount of fumbles (掉球率) he considers acceptable. These are constraints.

Types of Optimization Problems



- Some problems have constraints and some do not.

An example about Optimization problem with constrains.

$$\text{minimize } f(x),$$

$$\text{subject to } \; g(x) < 1$$
$$h(x) > 1$$

An example about Optimization problem without constrains

$$\text{minimize } f(x),$$

- Optimization problem with constrains is also called constrained optimization.
- Optimization problem without constrains is also called unconstrained optimization.

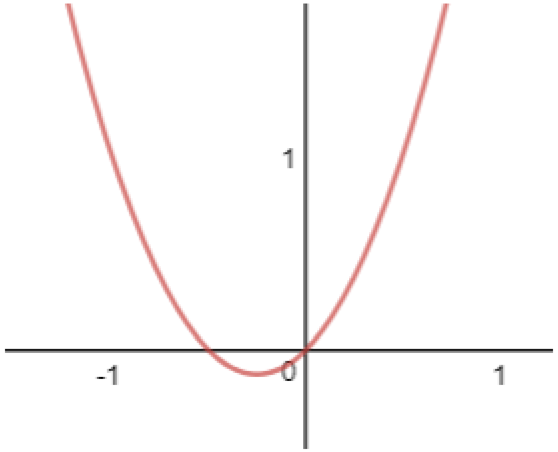
Credits from Dr. Julia Roberts and Dr. Mykel Kochenderfer

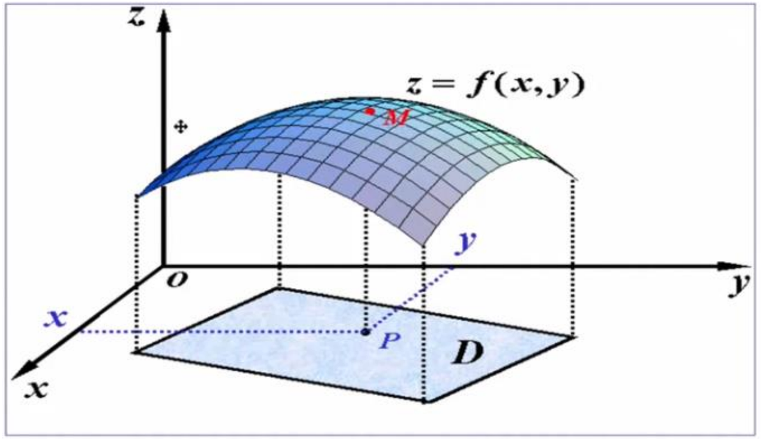
Types of Optimization Problems

- There can be one variable or many.
- For example: $x_1, x_2, x_3, \ldots, x_n \in R$,

  $f(x_1)$ only has one variable.

  $f(x_1, x_2, x_3, \ldots, x_n)$ has many variables.





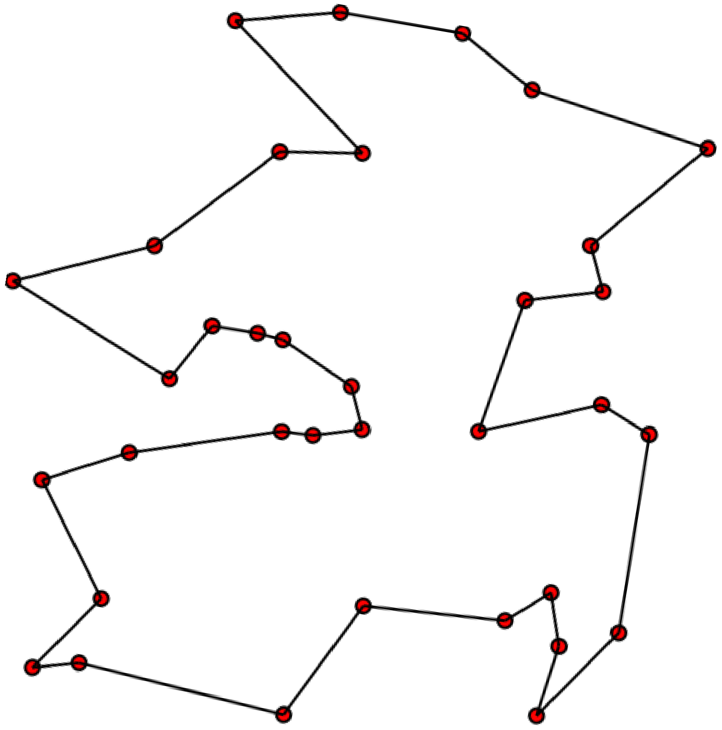
Credits from Dr. Julia Roberts and Dr. Mykel Kochenderfer

Types of Optimization Problems

- Variables can be discrete (for example, only have integer values) or continuous.

- Traveling Salesman Problem has discrete variables.

Given a list of cities and the distances between each pair of cities, what is the shortest possible route that visits each city exactly once and returns to the origin city?



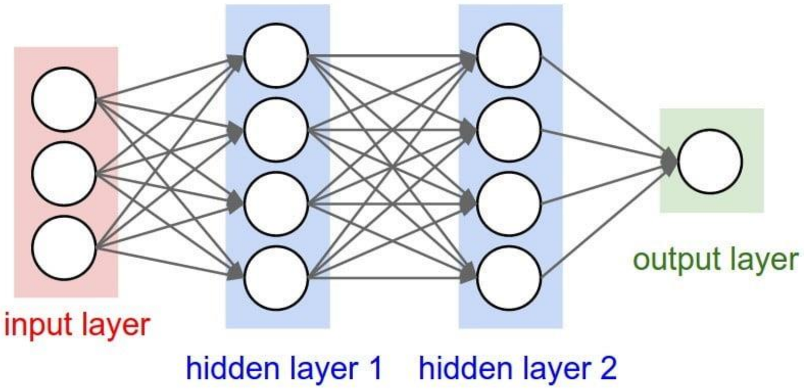
Because the number of possible routes are finite and discrete.

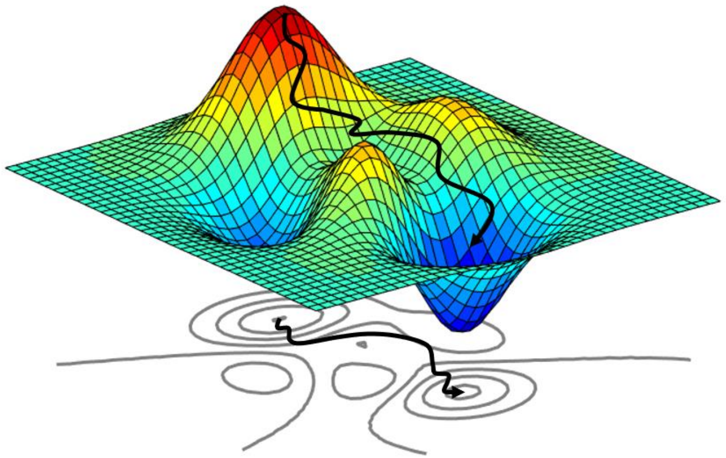
Credits from Wikpedia, Dr. Julia Roberts and Dr. Mykel Kochenderfer

Types of Optimization Problems

- Variables can be discrete (for example, only have integer values) or continuous.

- The problem optimizing the deep neural networks in machine learning has the continuous variables.





Credits from Dr. Julia Roberts and Dr. Mykel Kochenderfer

Types of Optimization Problems

- Some problems are static (do not change over time) while some are dynamic (continual adjustments must be made as changes occur).

- Traveling Salesman Problem is a static problem.

- Dynamic problem aims to find not just the maximum value of some function, but rather, the actual function that provides a time path for the values of the variables so that some value function is maximized or minimized over a given interval of time.

An example of dynamic problem:

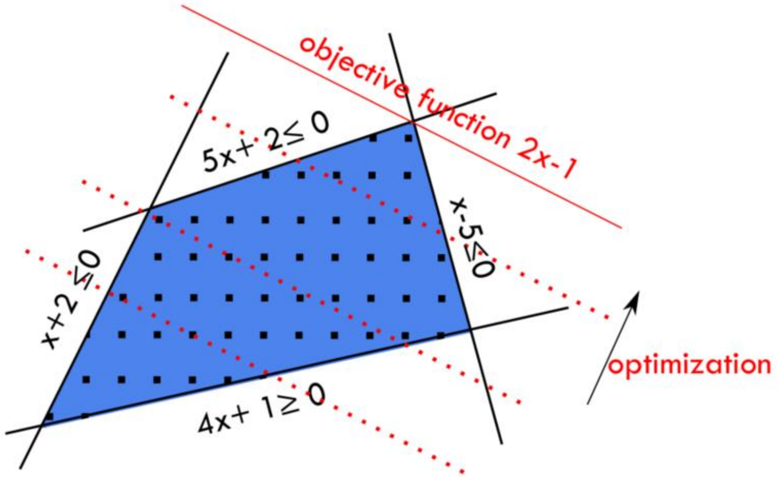
$$
\begin{array}{rll}
\min\limits_{u(t)} \left[M\left(\mathbf{x}\left(t_f\right), \mathbf{u}\left(t_f\right), t_f\right) + \int_{t_0}^{t_f} L\left(\mathbf{x}(t), \mathbf{u}(t), t\right) \mathrm{d}t \right] & \begin{array}{l}\text{optimization purpose} \\ (M \textit{ is the Mayer Term} \\ L \textit{ is the Lagrange Term}) \end{array} & (1) \\
\mathbf{0} = \mathbf{f}\left(\mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), t\right) & \begin{array}{l}\text{dynamical description} \\ \text{of the system}\end{array} & (2) \\
\mathbf{0} \leq \mathbf{g}\left(\mathbf{x}(t), \mathbf{u}(t), t\right) & \begin{array}{l}\text{path constraints} \\ \text{e.g. physical limits of components}\end{array} & (3) \\
\mathbf{0} = \mathbf{r}_0\left(\mathbf{x}\left(t_0\right), t_0\right) & \text{Initial constraint} & (4) \\
\mathbf{0} = \mathbf{r}_f(\mathbf{x}\left(t_f\right) \mathbf{u}\left(t_f\right), t_f) & \text{Final constraint} & (5)
\end{array}
$$

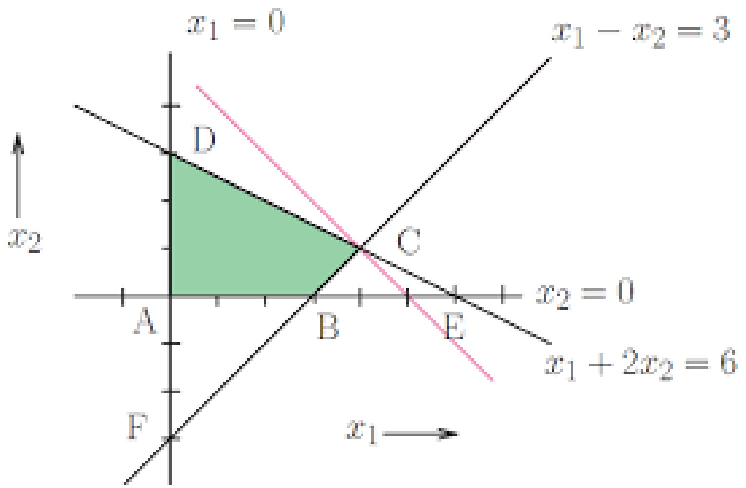
Credits from Dr. Julia Roberts and Dr. Mykel Kochenderfer

Types of Optimization Problems

- Equations can be linear (graph to lines) or nonlinear (graph to curves).

- Linear: linear optimization aims to achieve the best outcome (such as maximum profit or lowest cost) in a mathematical model whose requirements are represented by linear relationships.



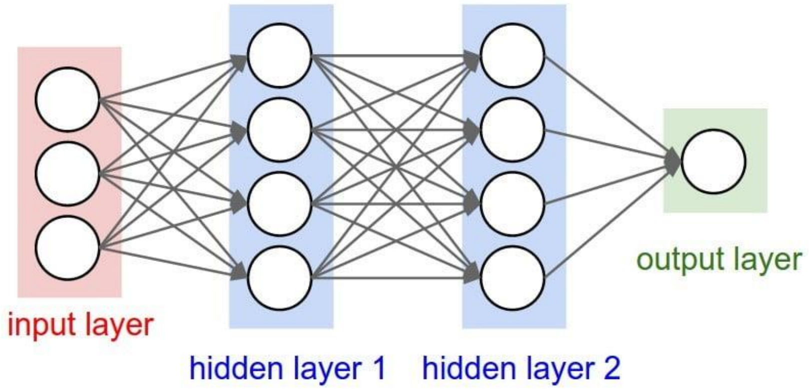


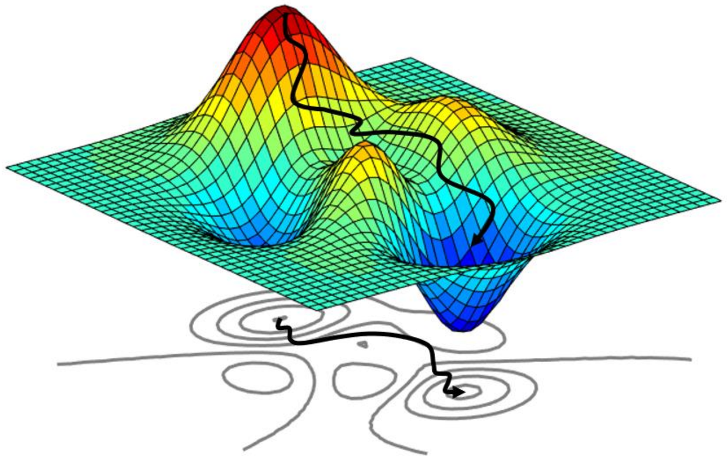
Credits from Wikpedia, Dr. Julia Roberts and Dr. Mykel Kochenderfer

Types of Optimization Problems

- Equations can be linear (graph to lines) or nonlinear (graph to curves).

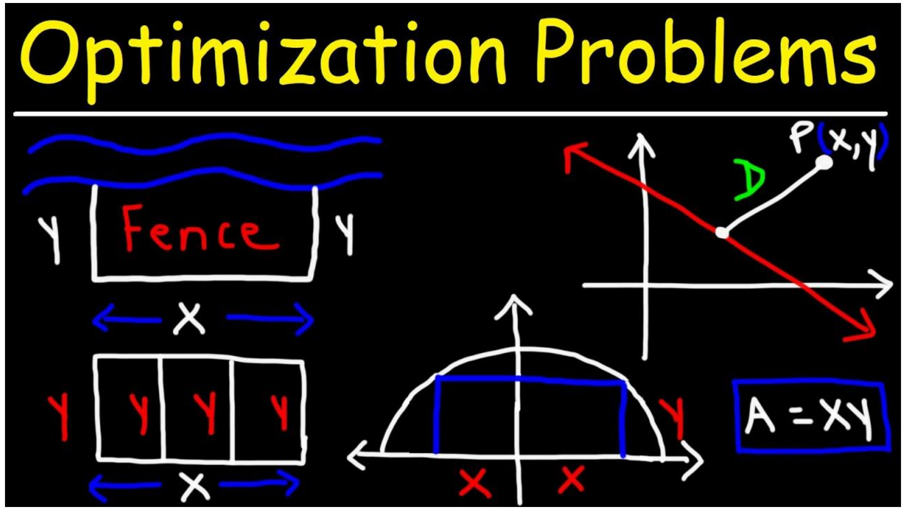
- Nonlinear : Nonlinear optimization means some of the constraints or the objective function are nonlinear.





Credits from Dr. Julia Roberts and Dr. Mykel Kochenderfer

Types of Optimization Problems

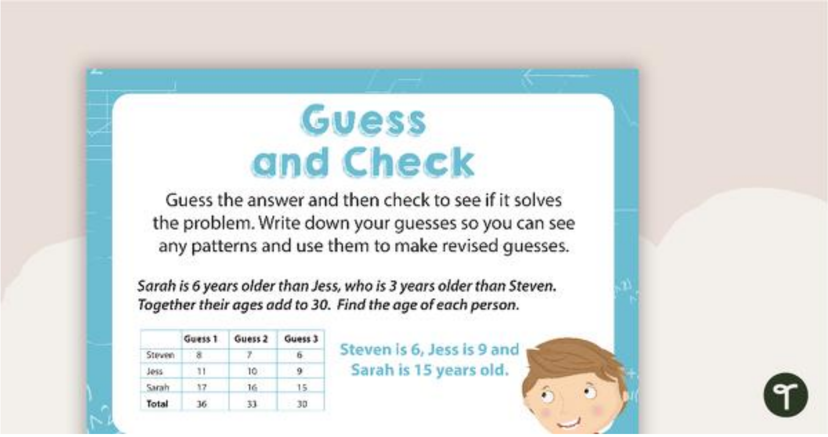


Some problems have constraints and some do not.

- There can be one variable or many.
- Variables can be discrete or continuous.
- Some problems are static (do not change over time) while some are dynamic (continual adjustments must be made as changes occur).
- Equations can be linear (graph to lines) or nonlinear (graph to curves)

Credits from Dr. Julia Roberts and Dr. Mykel Kochenderfer

Optimization Problems



Why Mathematical Optimization is Important?

- Mathematical Optimization works better than traditional "guess-and-check" methods

- Mathematical Optimization is a lot less expensive than traditional "guess-and-check" methods

- Optimization is at the heart of the Prescriptive stage indicating how to use resources efficiently to achieve the best possible goal under a series of conditions.

Credits from Dr. Julia Roberts and Dr. Mykel Kochenderfer

Optimization Problems: Summary

- Optimization problem in which the objective and constraints are given as mathematical functions and functional relationships.

Minimize $f(x_1, x_2, \ldots, x_n)$

Subject to

$h_i(x_1$

$x_2, \ldots, x_n) = \; a_i$, i=1,....,b

$g_j(x_1, x_2, \ldots, x_n) \leq \; b_j$, j=1,...,c

Exercises:

An airplane designer is trying to build the most fuel-efficient airplane possible, which is not slower than a speed c. Write one factor as an objective (“Minimize/maximize______”) and the rest as constraints (“______ ≤c1”, or ≥ or =). Delete any non-numerical factors:

- speed, fuel consumption, range, noise, weight, type of propulsion, cost, ease of use, amount of lift, amount of drag

Credits from Dr. Julia Roberts and Dr. Mykel Kochenderfer

Exercises:

Solution:

Because most fuel-efficient airplane, so Minimize fuel consumption

Because not slower than a speed c, so speed≥c

The answer is

Minimize: fuel consumption

Subject to speed≥c

Exercises:

When choosing a new phone and plan, you might consider:

- hours of talk time per month (not smaller than 3 hours);
- cost per month (not larger than 50 dollars);
- amount of storage (not smaller than 32 G);
- brands of available phones (the Apple or Huawei);
- cost of the phone (the cheapest);

Exercises:

Solution:

1. hours of talk time per month (not smaller than 3 hours);

   hours of talk time per month $\geq$ 3 hours

2. cost per month (not larger than 50 dollars);

   cost per month $\leq$ 50 dollars

3. amount of storage/memory (not smaller than 32 G)

   amount of storage$\geq$32 G

4. brands of available phones (the Apple or Huawei);

   brands of available phones = Apple or Huawei

5. cost of the phone (the cheapest)

   Minimize cost of the phone

Exercises:

The Solution is:

Minimize: cost of the phone

Subject to

1. hours of talk time per month $\geq$ 3 hours
2. cost per month $\leq$ 50 dollars
3. amount of storage $\geq$ 32 G
4. brands of available phones = Apple or Huawei

Exercises:

Consider a chocolate manufacturing company that produces only two types of chocolate – A and B. Both the chocolates require Milk and Choco only. To manufacture each unit of A and B, the following quantities are required:

- Each unit of A requires 1 unit of Milk and 3 units of Choco
- Each unit of B requires 1 unit of Milk and 2 units of Choco

The company kitchen has a total of 5 units of Milk and 12 units of Choco. On each sale, the company makes a profit of 6 dollors per unit A sold and 5 dollors per unit B sold.

Now, the company wishes to maximize its profit. How units of A and B should it produce respectively?

Exercises:

Step 1: Identify the variables.

the total number of units produced by A be = x

the total number of units produced by B be = y

Step 2: Formulate the objective function. Check whether the function needs to be minimized or maximized.

The total profit the company makes is given by the total number of units of A and B produced multiplied by its per-unit profit of 6 dollors and 5 dollors respectively.

Profit: Maximize $6x+5y$

Exercises:

Step 3: Write down the constraints.

- The company kitchen has a total of 5 units of Milk and 12 units of Choco.

- Each unit of A and B requires 1 unit of Milk.

  So $x+y\leq 5$

- Each unit of A and B requires 3 units & 2 units of Choco respectively.

  So $3x+2y\leq 12$

Exercises:

Step 3: Write down the constraints.

Also, the values for units of A can only be integers.

So we have two more constraints, $x \geq 0$ & $y \geq 0$.

Solution: $\underset{x,y \in Z}{\text{Maximize}}\ 6x + 5y$

Subject to $x+y \leq 5$;

$3x+2y \leq 12$;

$x \geq 0$;

$y \geq 0$,

where Z is the set consisting of all integers.

Exercises:

Solution: $\underset{x,y \in Z}{\text{Maximize}}\ 6x + 5y$

Subject to $x+y \leq 5;$

$3x+2y \leq 12;$

$x \geq 0;$

$y \geq 0.$

Above issue is related to Linear Optimization.

Next, we discuss linear optimization issue detaily.

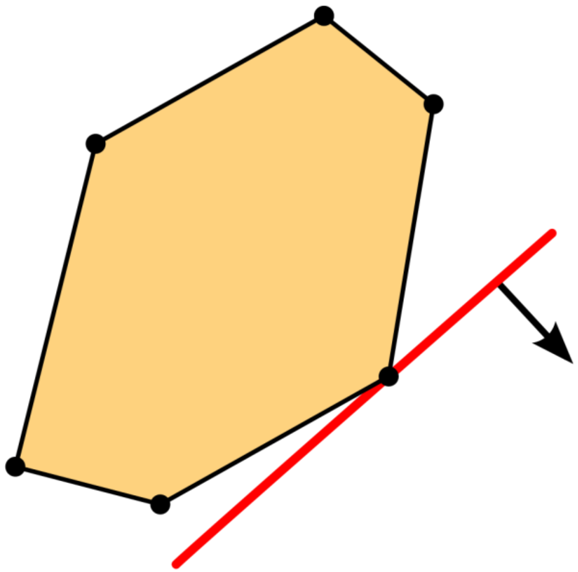
Linear Optimization (LO):

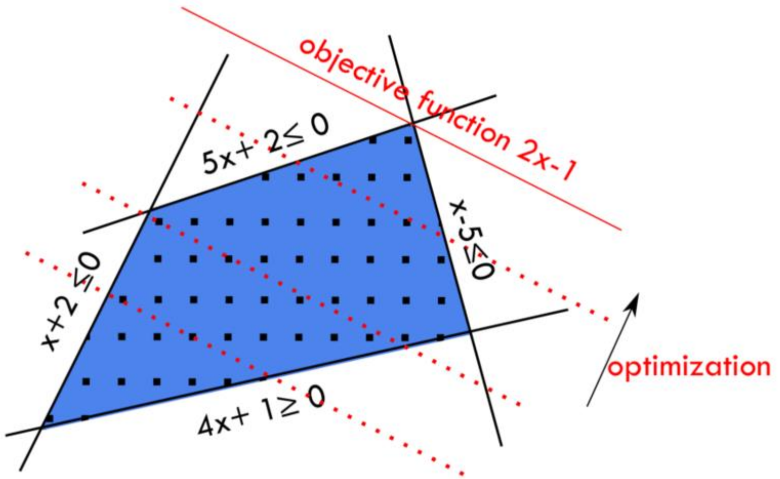
- Linear – all the functions are linear:

$$f(x_1,x_2,\ldots,x_n) = \sum_{i=1}^{n} a_i * x_i + a_0$$

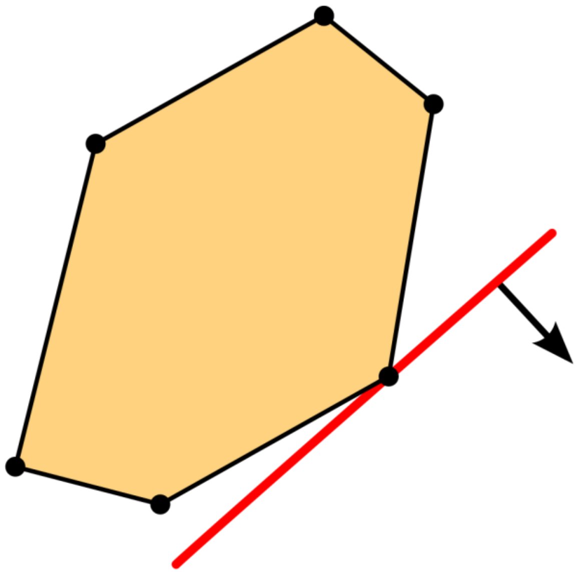
$$h(x_1,x_2,\ldots,x_n) = \sum_{i=1}^{n} b_i * x_i + b_0$$

$$g(x_1,x_2,\ldots,x_n) = \sum_{i=1}^{n} c_i * x_i + c_0$$





Linear Optimization (LO):



- A linear optimization model consists of:

1. Variables x1, x2,...,xn

2. A (linear) objective function:
   $f(x_1,x_2,\ldots,x_n)= \sum_{i=1}^{n} a_i * x_i + a_0$

3. A set of (linear) constraints:
$h_1, h_2,\ldots,h_d; g_1, g_2,\ldots, g_c$ (see Page 16)

Linear Optimization (LO): Examples

1. Objective Function: $f(x,y) = ax + by$;
2. Constraints: $h_1(x,y) = lx+my = h$;

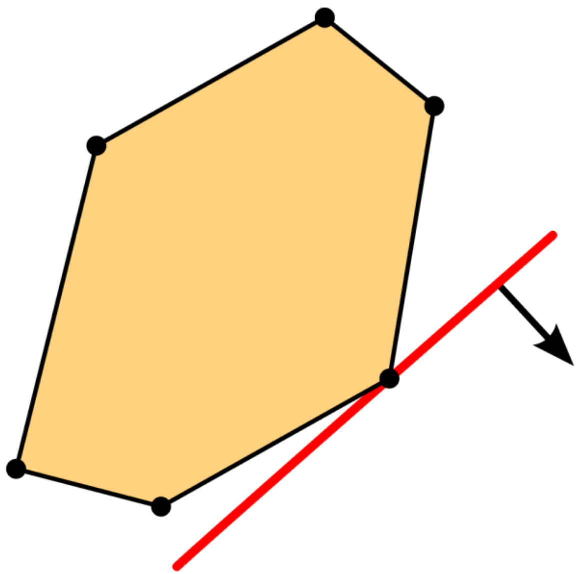
   $g_1(x,y) = cx + dy \le e$, $g_2(x,y) = fx + gy \le h$.

- Minimize $f(x,y) = 2x + 2y$;

  Subject to $h_1(x,y) = 3x+4y = 0$;

  $g_1(x,y) = 5x + 2y \le 0$;

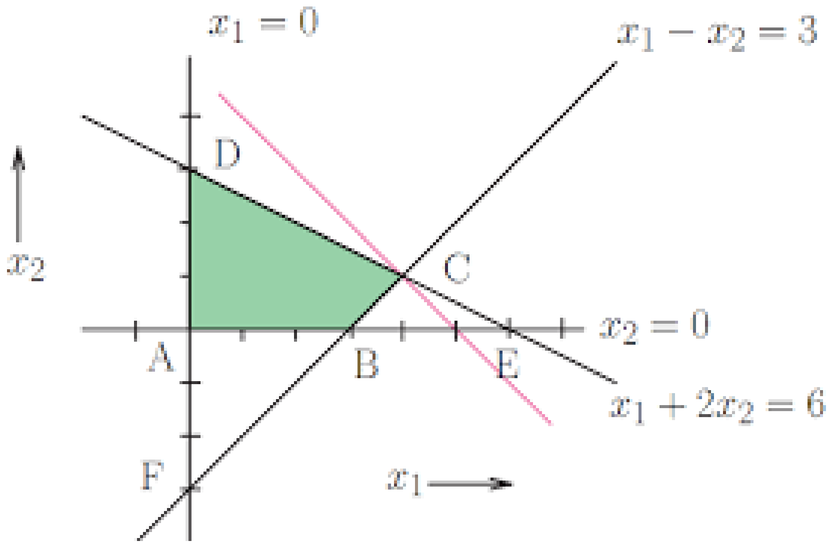
  $g_2(x,y) = 8x + 9y \le 10$.



We will discuss how to address this issue in page 43.

Linear Optimization (LO):

Advantages of Linear Optimization:



- The main advantage of LO is
its simplicity and easy way to understand.

- LO makes use of available resources.

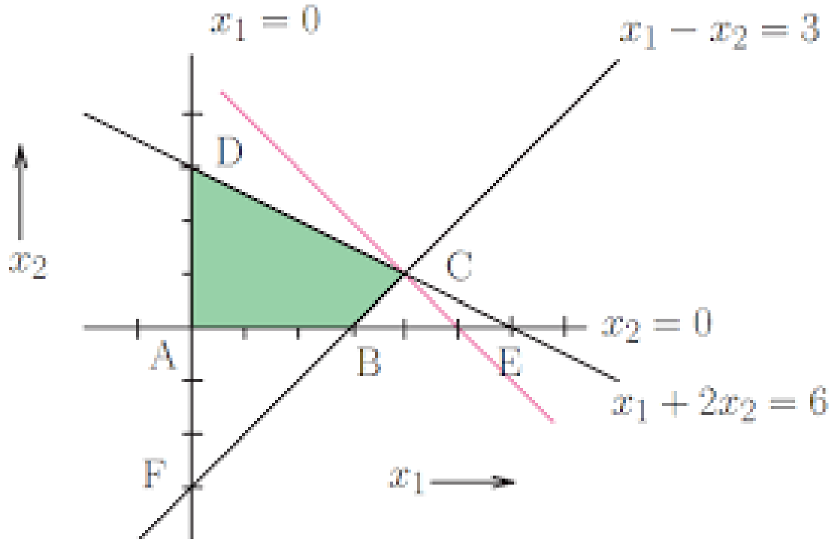
- To address many diverse combination problem.

- LO is adaptive and more flexibility to analyze the problem.

Credits from Dr. Amulya Bharadwaj and Mr. Mahadeva Prasad

Linear Optimization (LO):

Disadvantages of Linear Optimization:



- LO works only for linear system.
- LO addresses static problem.
  It doesn't consider change and evolution of variables.

Credits from Dr. Amulya Bharadwaj and Mr. Mahadeva Prasad

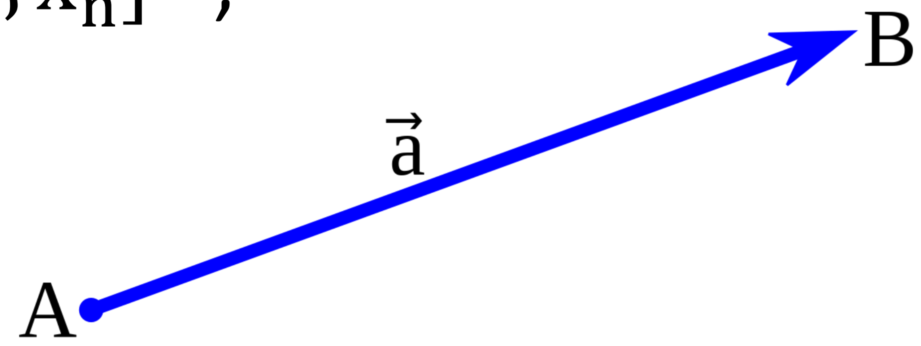
Matrix Representation of LO

To represent a Linear Optimization problem more convenient, we use matrix to rewritten it.

First, the variables are $x_1, x_2, \ldots, x_n$, we can use a vector to represent it:

$$\mathbf{x} = [1, x_1, x_2, \ldots, x_n]^T,$$

where T is the Transpose of vector.



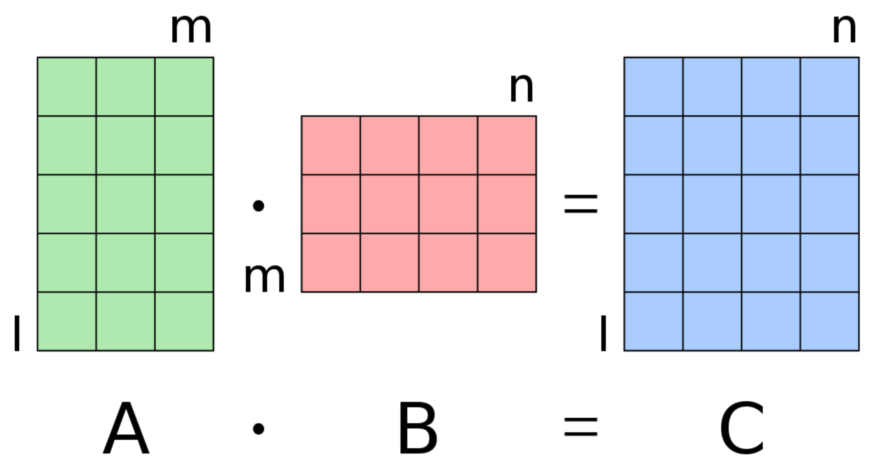
Matrix Representation of LO

Second, the objective function

$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} a_i * x_i + a_0$

can be written as $f(\mathbf{x}) = a\mathbf{x}$,

where $a = [a_0, a_1, a_2, \ldots, a_n]$

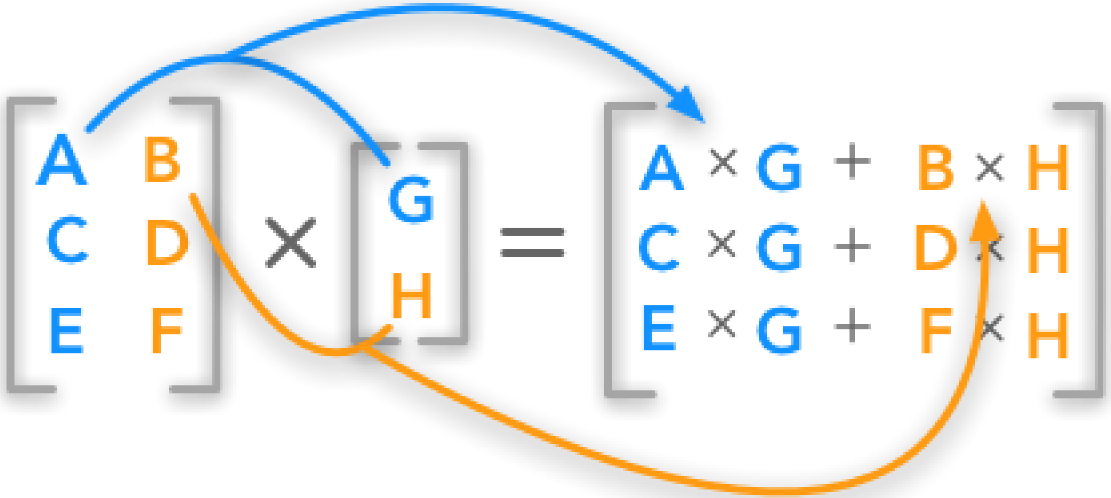


Matrix Representation of LO

Third, the constraints can be written as

$$\mathbf{h}(\mathbf{x}) = \mathbf{B}\mathbf{x} = \mathbf{0}, \quad \mathbf{g}(\mathbf{x}) = \mathbf{C}\mathbf{x} \leq \mathbf{0},$$

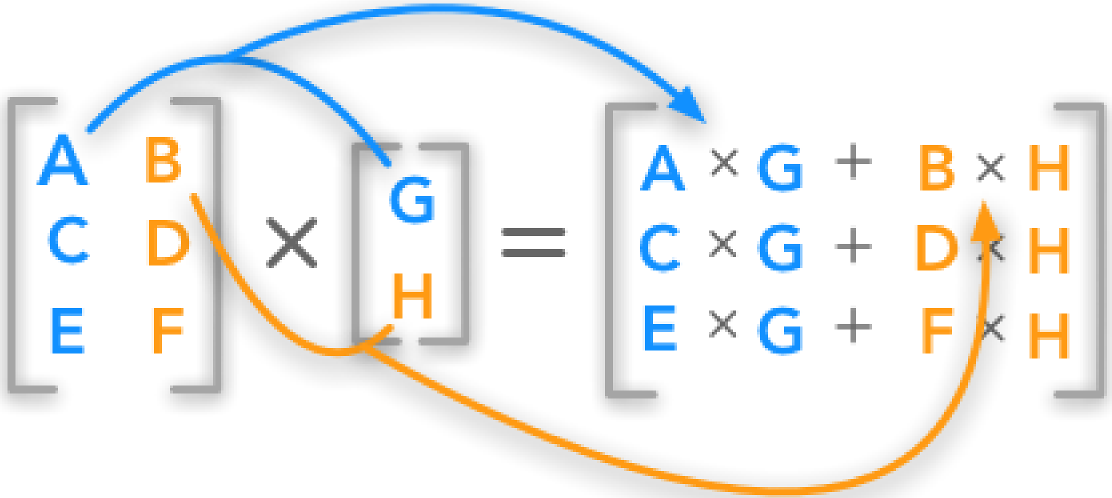
where B is a dx(n+1) matrix and C is a cx(n+1) matrix, here d is the number of constraints with respect to **h,** and c is the number of constraints with respect to **g**.



Matrix Representation of LO

$$\mathbf{Cx} \leq \mathbf{0}$$

means that all elements in vector Cx are not larger than 0.

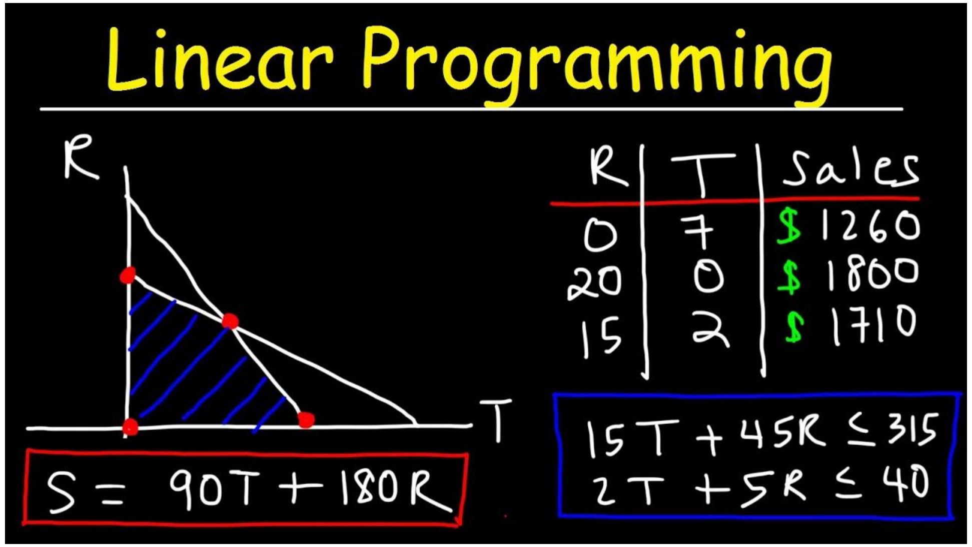


Matrix Representation of LO

Lastly, the LO problem can be rewritten as follows:

$$\begin{aligned} &\text{Minimize } a\mathbf{x} \\ &\text{Subject to } \mathbf{Bx = 0;} \\ &\mathbf{Cx} \le \mathbf{0}. \end{aligned}$$

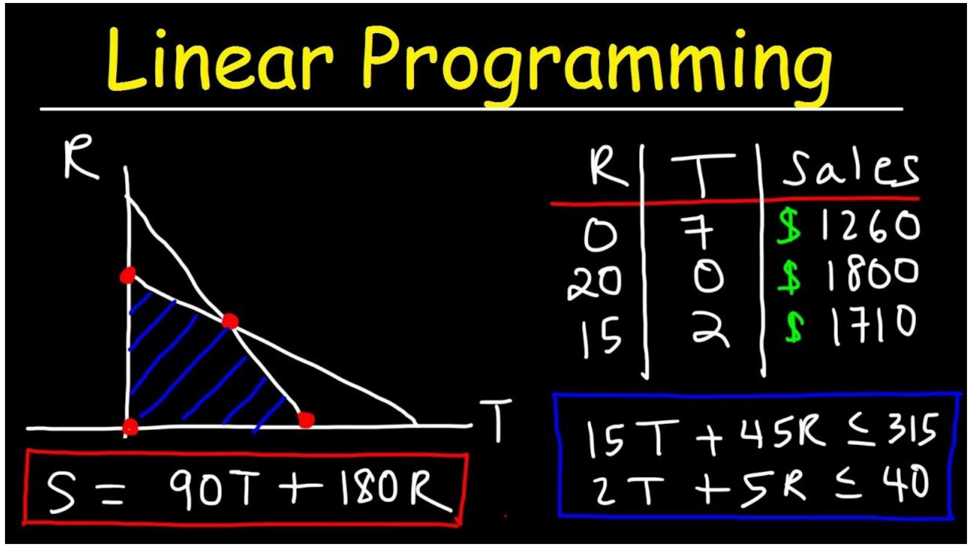
We will give examples to help understnad.



Matrix Representation of LO: Exercises

Please give the Matrix Representation of the following LO problem:

- Minimize $f(x,y) = 2x + 2y$;
  Subject to $h1(x,y) = 3x+4y = 0$;
  $g1(x,y) = 5x + 2y \leq 0$;
  $g2(x,y) = 8x + 9y \leq 10$.



Matrix Representation of LO: Exercises

Please give the Matrix Representation of the following LO problem:

Let **x** = $[1, \mathrm{x}, \mathrm{y}]^{\mathrm{T}}$

First, f(x,y) = 2x+2y = $[0 \quad 2 \quad 2]$**x**;

Second, h1(x,y) = 3x+4y =$[0 \quad 3 \quad 4]$**x**;

Third, g1(x,y) = 5x + 2y ≤ 0; g2(x,y) = 8x + 9y ≤ 10 can be written as

g(x) = $\begin{bmatrix} 0 & 5 & 2 \\ -10 & 8 & 9 \end{bmatrix}$**x**.

Matrix Representation of LO: Exercises

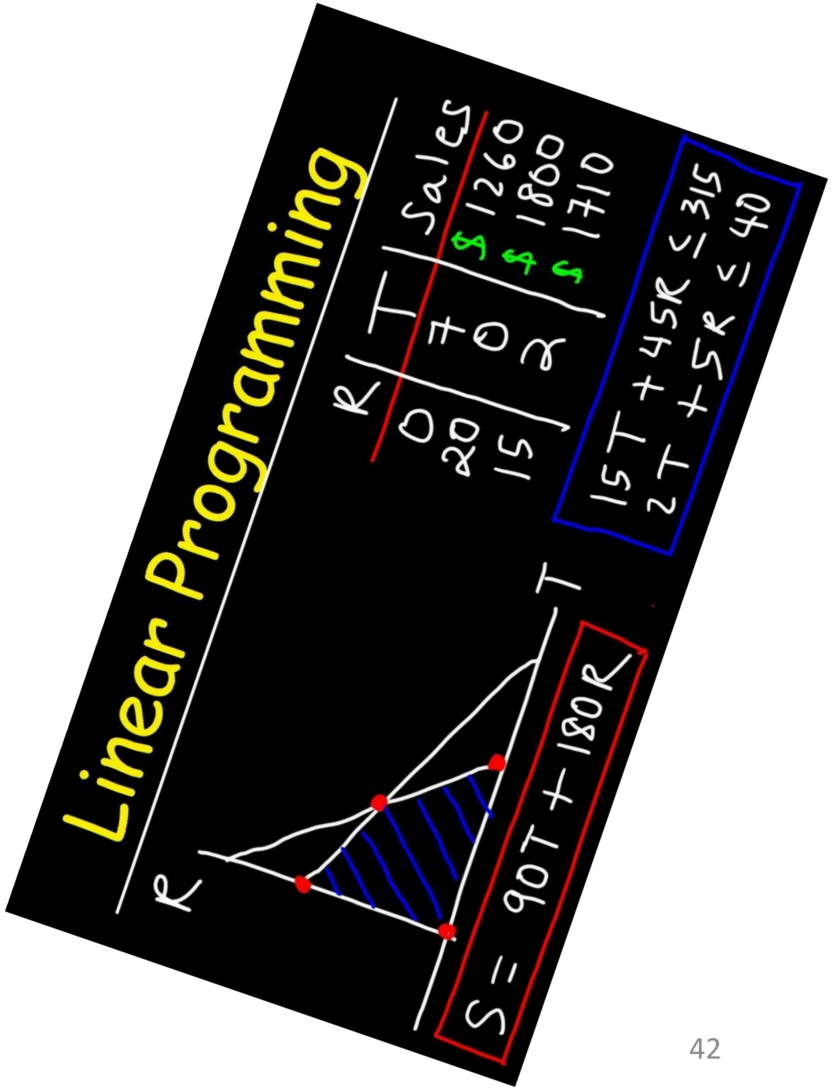
So the solution is

Minimize $[0 \quad 2 \quad 2]\mathbf{x}$

Subject to

$$[0 \quad 2 \quad 2]\mathbf{x=0}$$

$$\begin{bmatrix} 0 & 5 & 2 \\ -10 & 8 & 9 \end{bmatrix}\mathbf{x} \leq \mathbf{0}$$



Matrix Representation of LO: Exercises

Please give the Matrix Representation of the following LO problem:

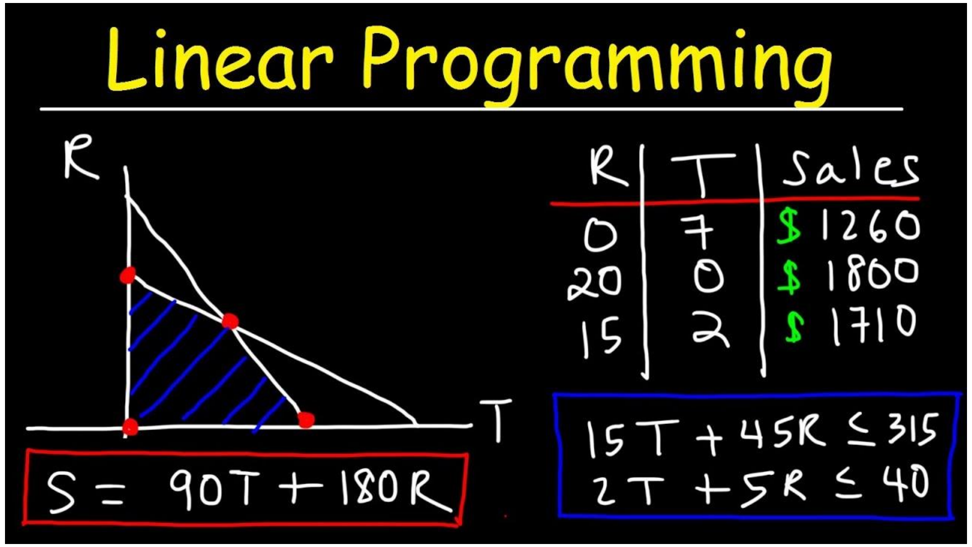
Maximum $6x+5y$

Subject to $x+y \leq 5;$

$3x+2y \leq 12;$

$x \geq 0;$

$y \geq 0.$



Matrix Representation of LO: Exercises

Please give the Matrix Representation of the following LO problem:

Let **x** = $[1, x, y]^T$

First, f(x,y) = 6x+5y = $[0 \quad 6 \quad 5]$**x**;

Second, x+y≤ 5;

3x+2y≤ 12;

x ≥ 0;

y ≥ 0.

can be rewritten as

$$\mathbf{g(x)} = \begin{bmatrix} -5 & 1 & 1 \\ -12 & 3 & 2 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}\mathbf{x}.$$

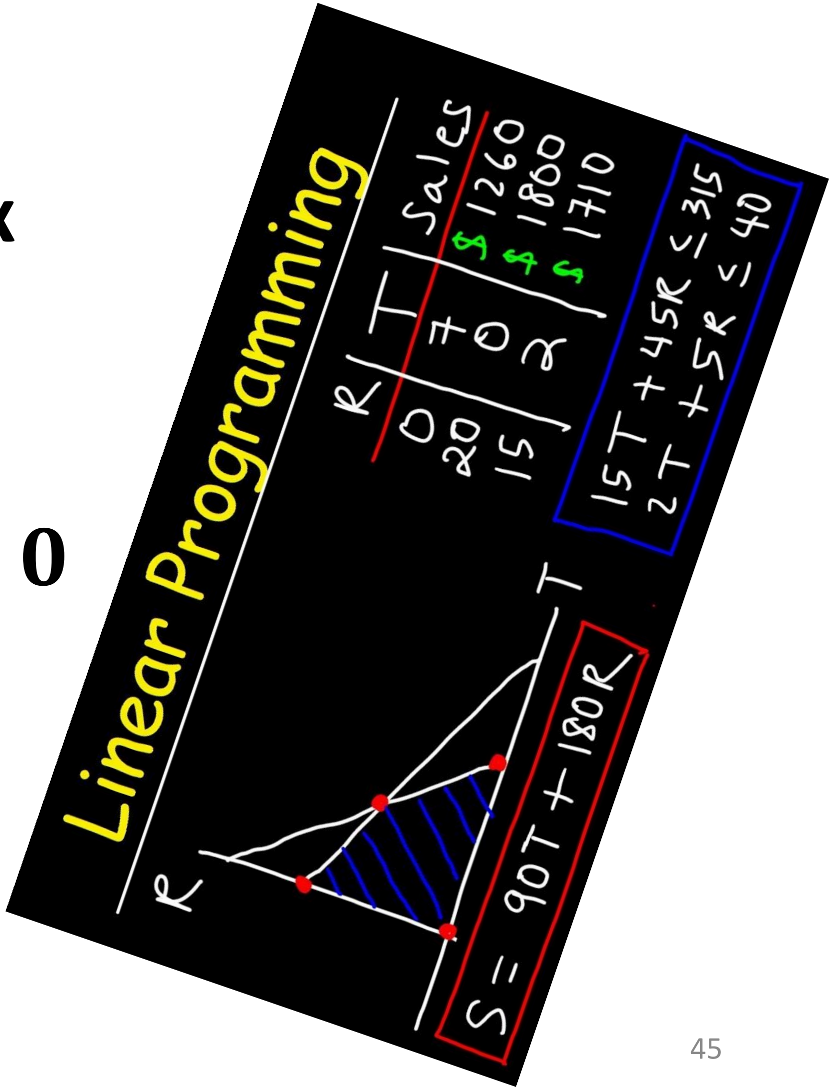
Matrix Representation of LO: Exercises

So the solution is

$$\text{Minimize } [0 \quad 2 \quad 2]\mathbf{x}$$

Subject to

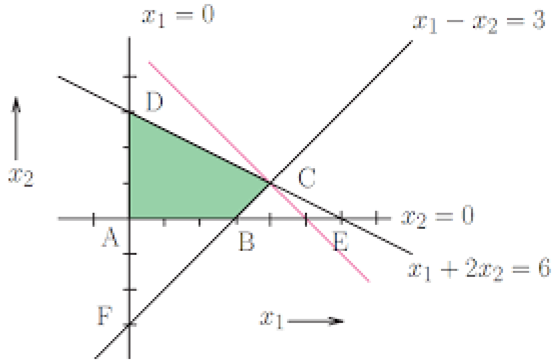
$$\begin{bmatrix} -5 & 1 & 1 \\ -12 & 3 & 2 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}\mathbf{x} \leq \mathbf{0}$$

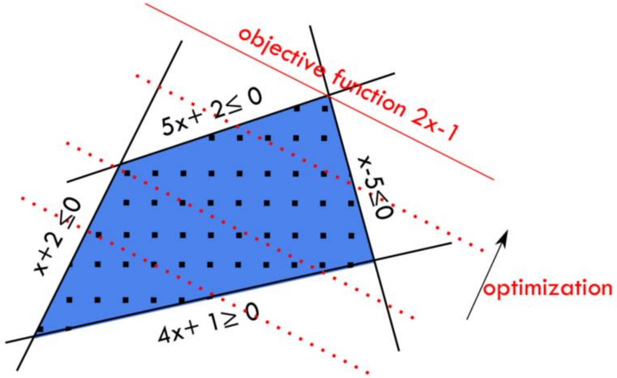


How to Address Linear Optimization

- Many types of algorithms have been developed over the years to solve them. Some famous mentions include the Simplex method, the Hungarian approach, and others. Here we are going to concentrate on one of the most basic methods to handle a linear programming problem i.e. the graphical method.

- Graphical Method works for almost all different types of problems but gets more and more difficult to solve when the number of decision variables and the constraints increases. Therefore, we'll illustrate it in a simple case i.e. for two variables only.





Graphical Method to Address Linear Optimization

- We will first discuss the steps of the algorithm:

Step 1: Construct a graph and plot the constraint lines

We have already understood the mathematical formulation of an LP problem in a previous section. Note that this is the most crucial step as all the subsequent steps depend on our analysis here.

We use following problem as an example:

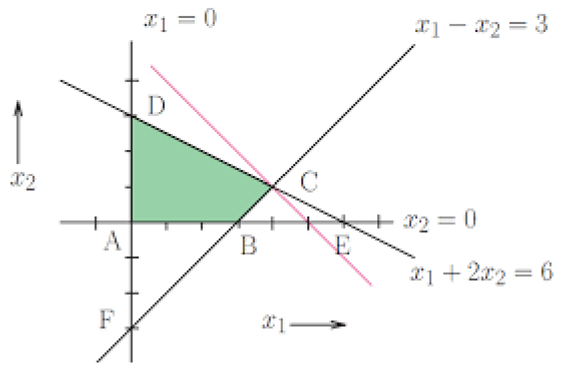
Maximum $6x_1+5x_2$

Subject to $x_1-x_2 \leq 3;$

$x_1+2x_2 \leq 6;$

$x_1 \geq 0;$

$x_2 \geq 0.$



Credits from topper.com

Graphical Method to Address Linear Optimization

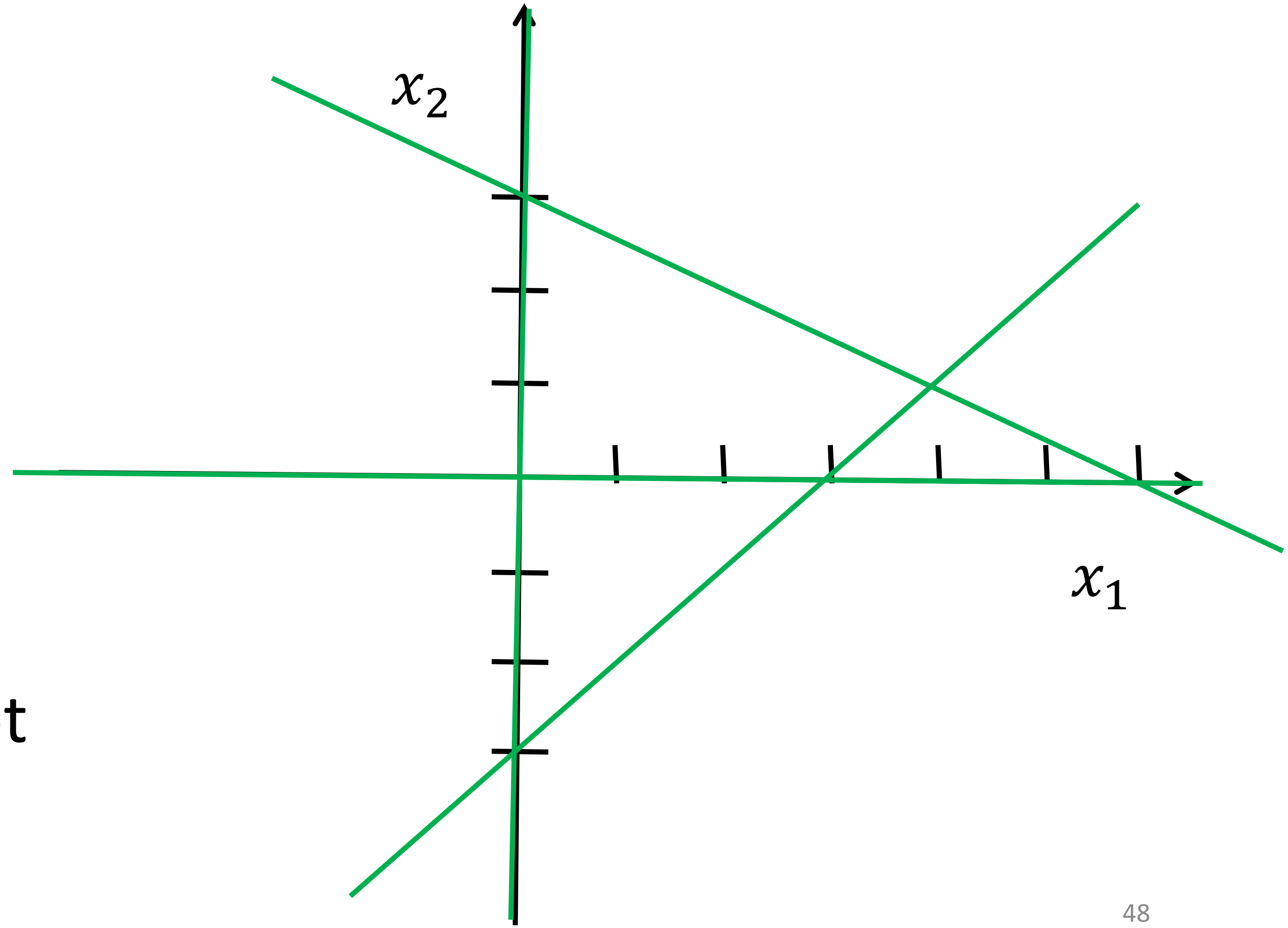
Maximum $6x_1+5x_2$

Subject to $x_1-x_2 \leq 3;$

$x_1+2x_2 \leq 6;$

$x_1 \geq 0;$

$x_2 \geq 0.$



Construct a graph and plot the constraint lines

Graphical Method to Address Linear Optimization

Step 2: Determine the valid side of each constraint line

This is used to determine the domain of the available space, which can result in a feasible solution.

How to check? A simple method is to put the coordinates of the origin (0,0) in the problem and determine whether the objective function takes on a physical solution or not. If yes, then the side of the constraint lines on which the origin lies is the valid side. Otherwise it lies on the opposite one.

Credits from topper.com

Graphical Method to Address Linear Optimization

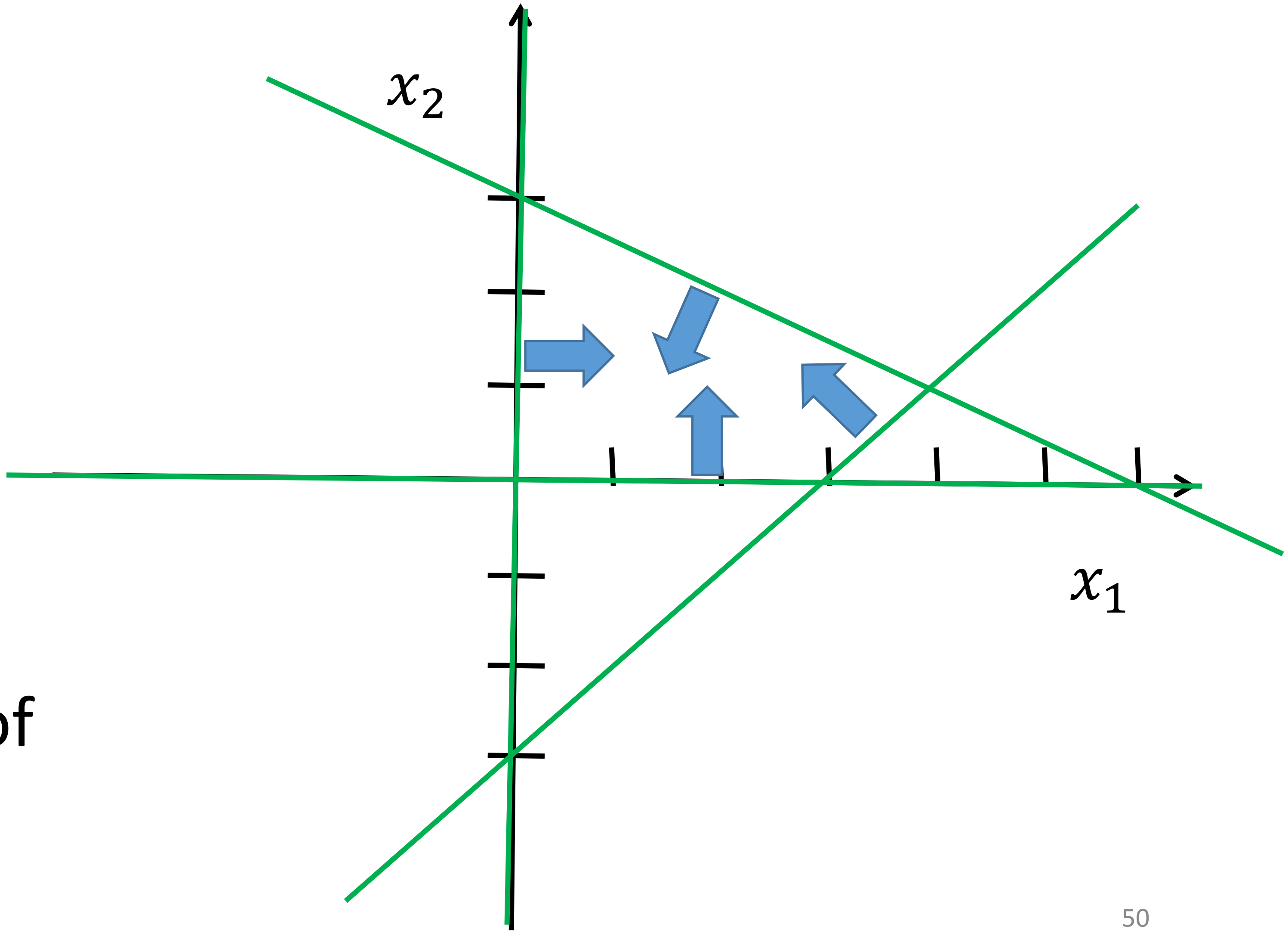
Maximum $6x_1+5x_2$

Subject to $x_1-x_2 \leq 3;$

$x_1+2x_2 \leq 6;$

$x_1 \geq 0;$

$x_2 \geq 0.$



Determine the valid side of each constraint line

Graphical Method to Address Linear Optimization

Step 3: Identify the feasible solution region

The feasible solution region on the graph is the one which is satisfied by all the constraints. It could be viewed as the intersection of the valid regions of each constraint line as well.

Choosing any point in this area would result in a valid solution for our objective function.

Graphical Method to Address Linear Optimization

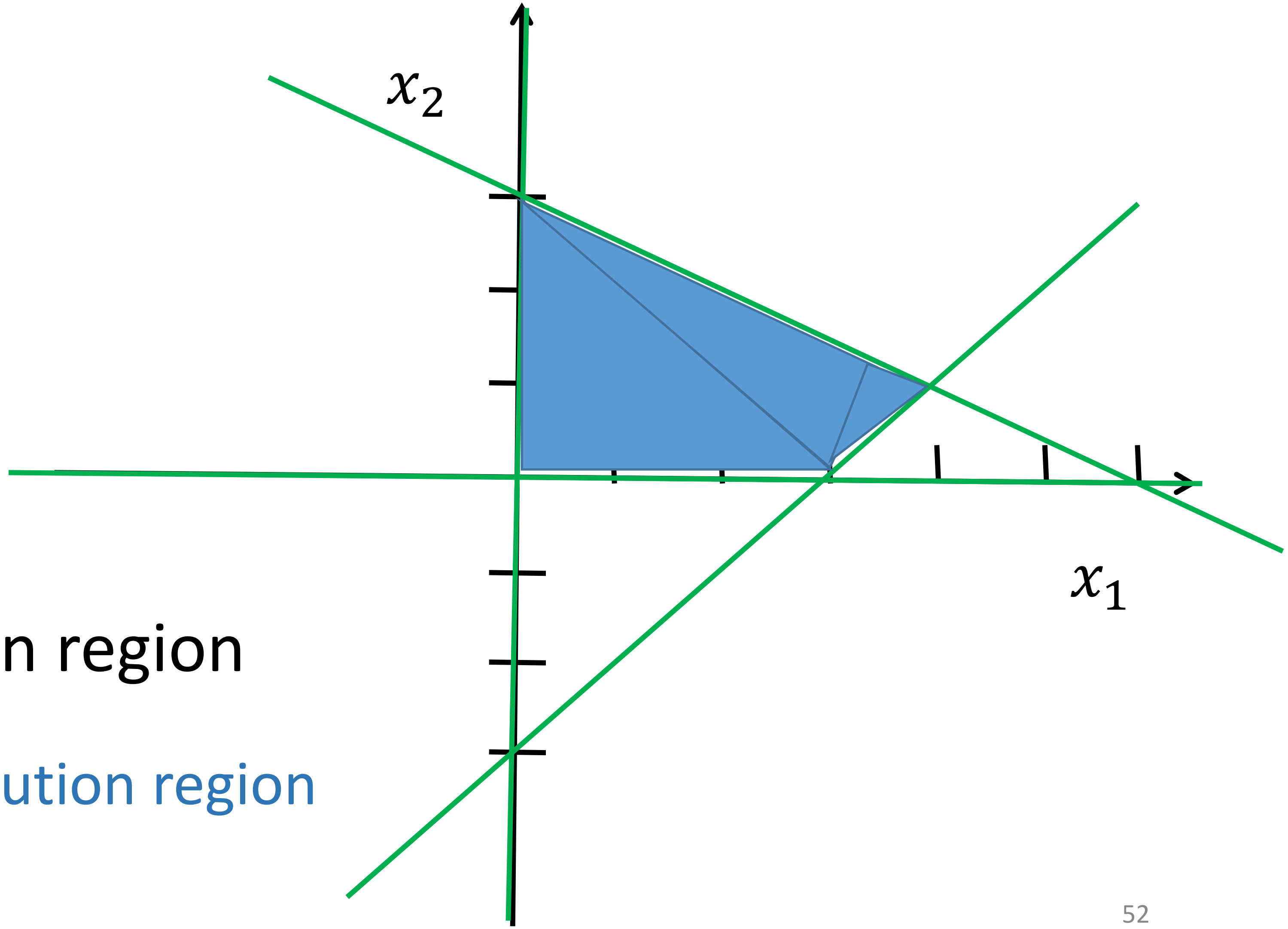
Maximum $6x_1+5x_2$

Subject to $x_1\text{-}x_2 \leq 3;$

$x_1+2x_2 \leq 6;$

$x_1 \geq 0;$

$x_2 \geq 0.$



Identify the feasible solution region

The blue part is the feasible solution region

Graphical Method to Address Linear Optimization

Step 4: Plot the objective function on the graph

It will clearly be a straight line since we are dealing with linear equations here. One must be sure to draw it differently from the constraint lines to avoid confusion. Choose the constant value in the equation of the objective function randomly, just to make it clearly distinguishable.

Graphical Method to Address Linear Optimization

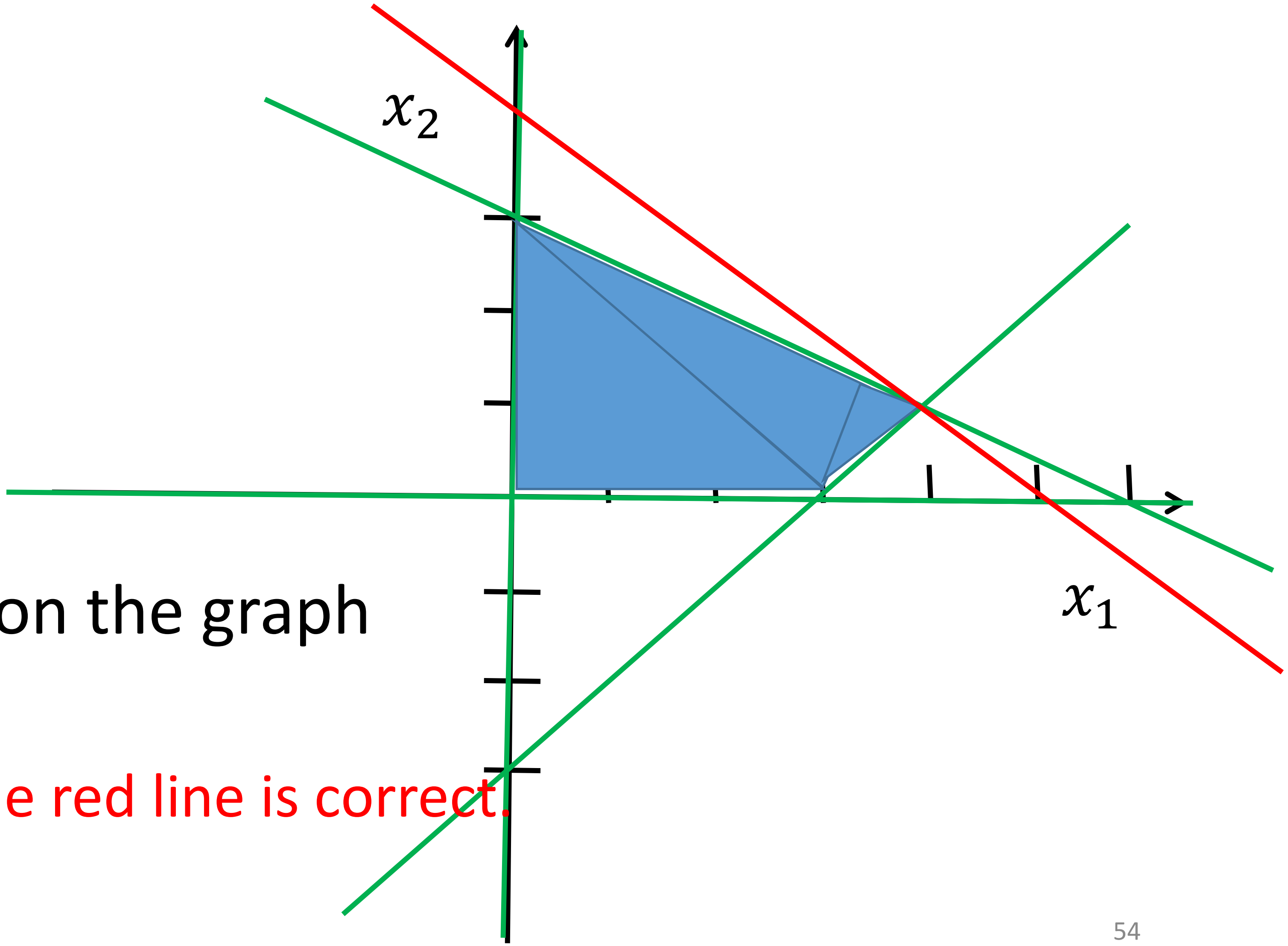
Maximum $6x_1+5x_2$

Subject to $x_1-x_2 \leq 3;$

$x_1+2x_2 \leq 6;$

$x_1 \geq 0;$

$x_2 \geq 0.$



Plot the objective function on the graph

The red line

In next page, we discuss why the red line is correct.

Graphical Method to Address Linear Optimization

Maximum $6x_1$+$5x_2$

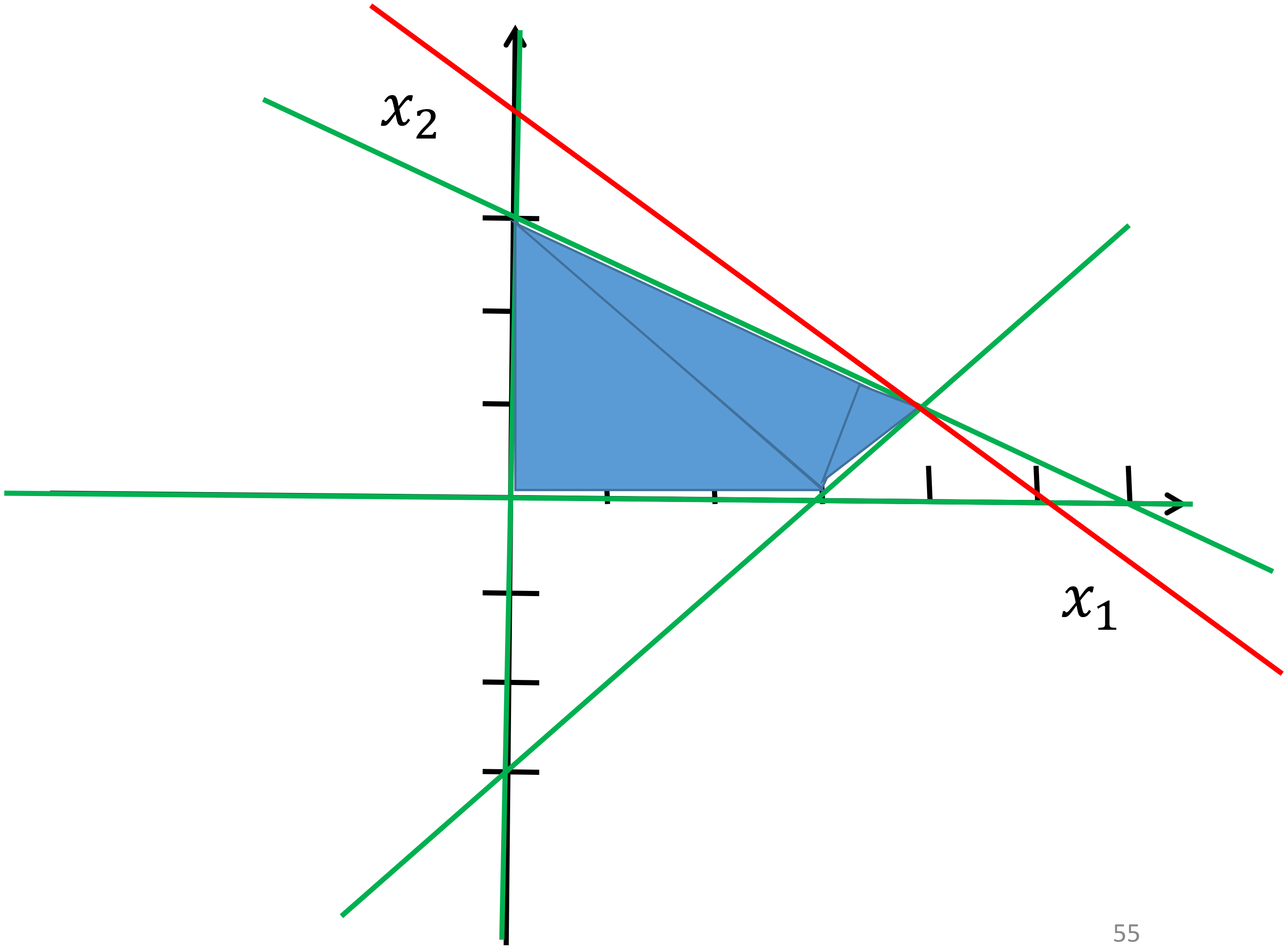
Subject to x_1-x_2 ≤ 3;

x_1+$2x_2$ ≤ 6;

$x_1 \geq 0$;

$x_2 \geq 0$.

Why is the red line correct?



For a line,

$6x_1 + 5x_2 = k$.

k/5 is the point of intersection bewteen

$6x_1 + 5x_2 = k$ and y-axis

Graphical Method to Address Linear Optimization

Maximum $6x_1$+$5x_2$

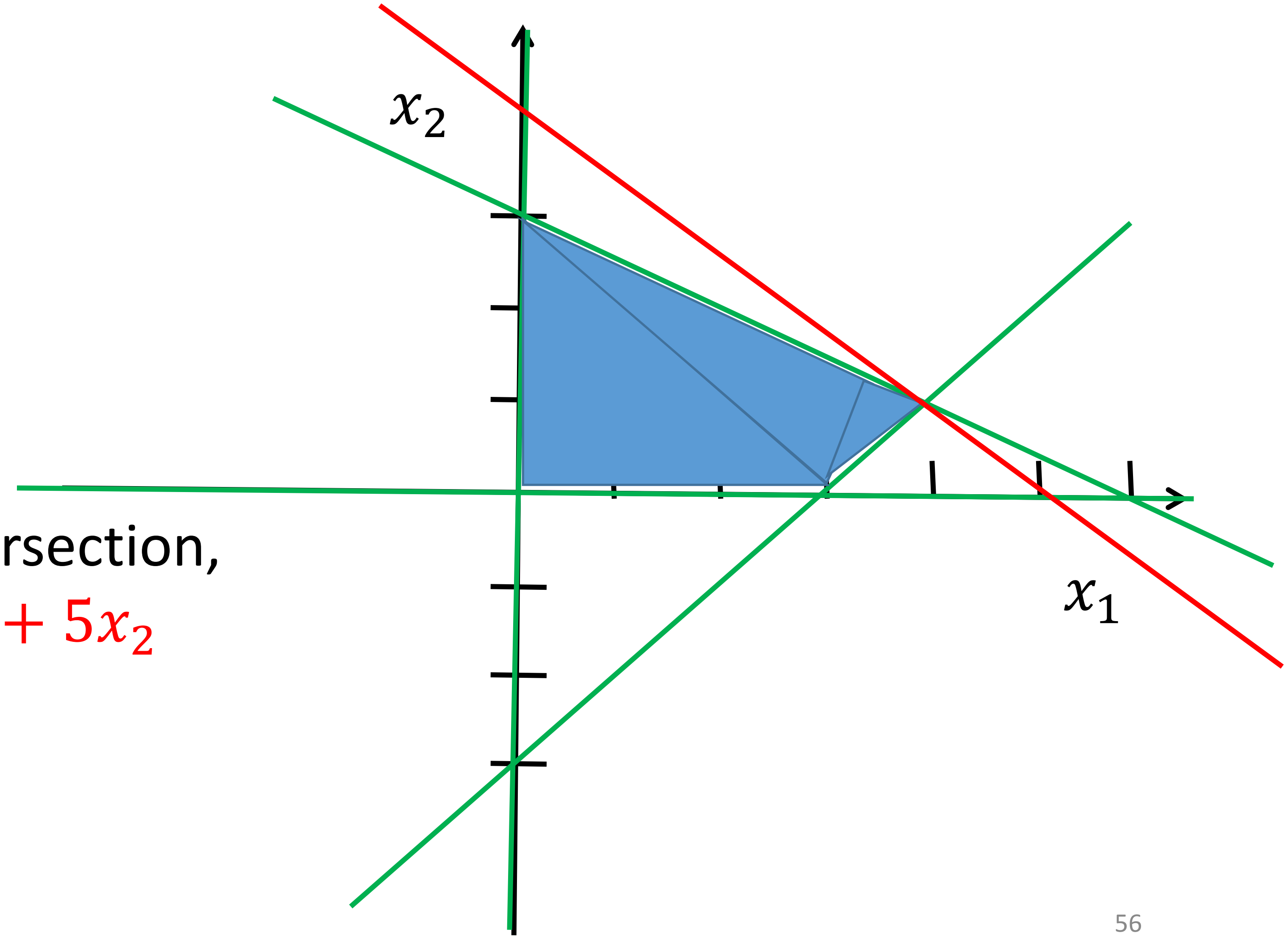
Subject to x_1-$x_2 \leq 3$;

x_1+$2x_2 \leq 6$;

$x_1 \geq 0$;

$x_2 \geq 0$.

Why is the red line correct?



So the larger the point of intersection, the larger the value that $6x_1 + 5x_2$ takes.

That is why we select the red line.

Graphical Method to Address Linear Optimization

Question: If we replace maximum by minimize, then what the red line is?

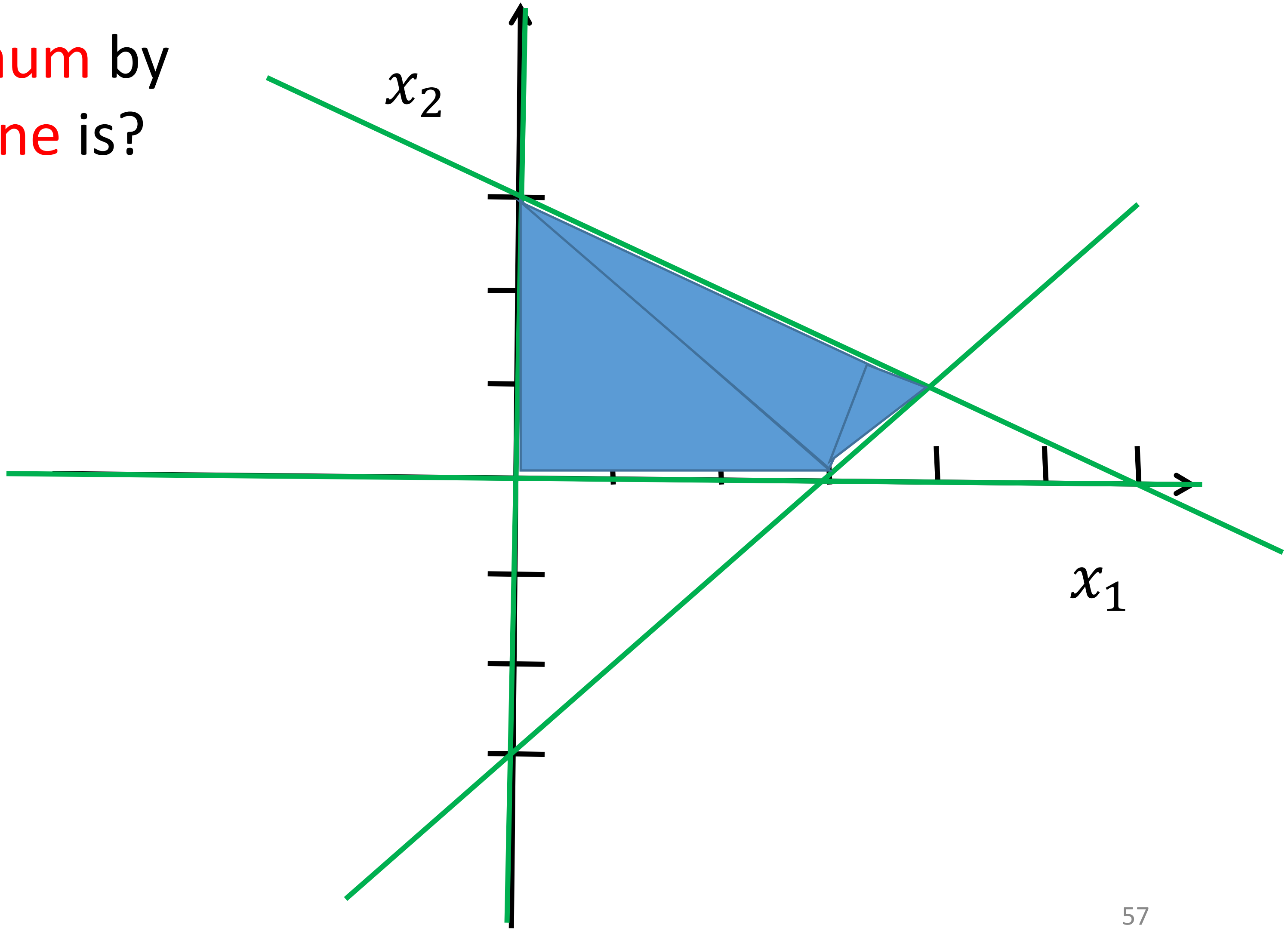
Minmize $6x_1+5x_2$

Subject to $x_1-x_2 \leq 3;$

$x_1+2x_2 \leq 6;$

$x_1 \geq 0;$

$x_2 \geq 0.$



Graphical Method to Address Linear Optimization

Question: If we replace maximum by minimize, then what the red line is?

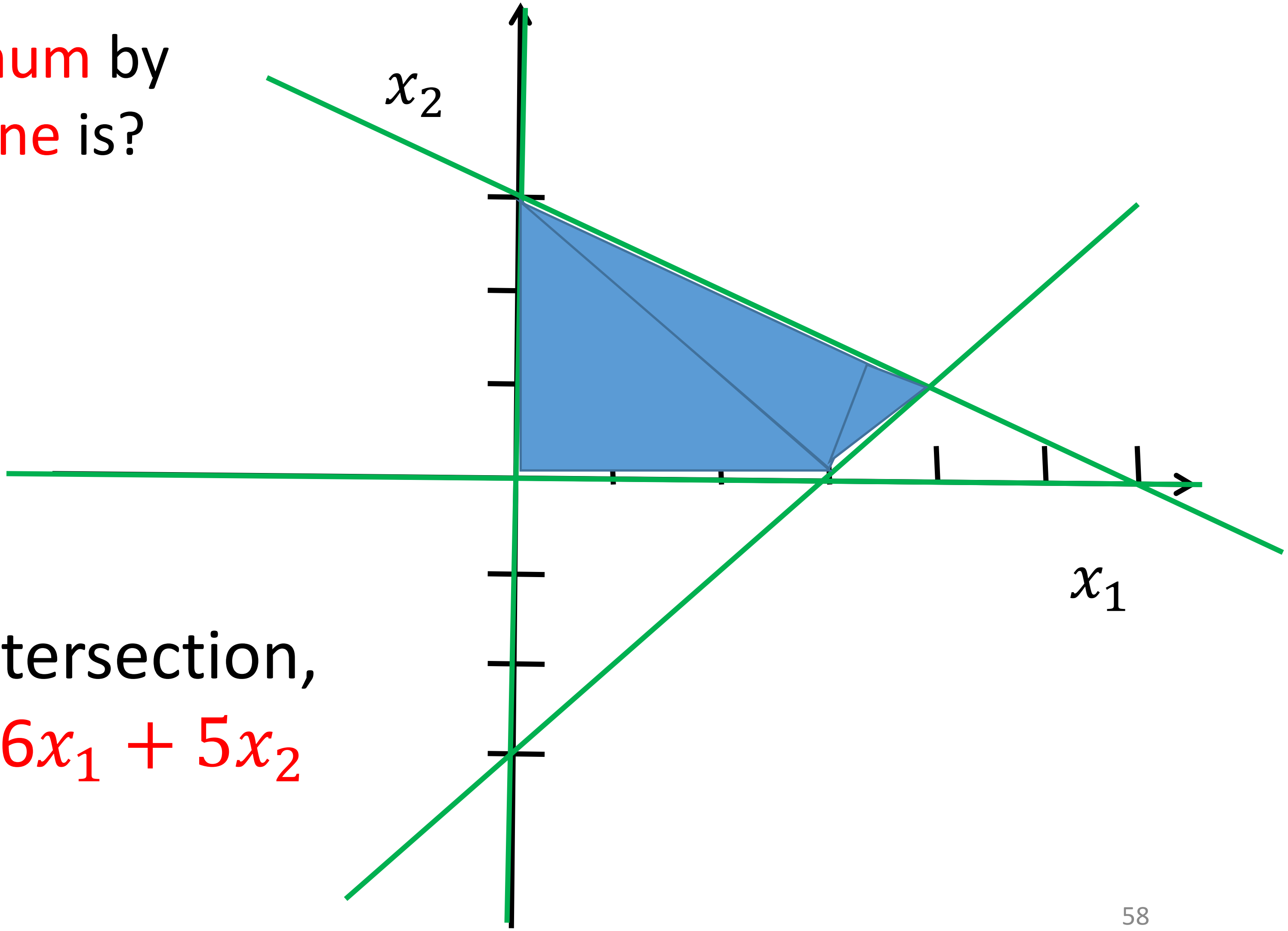
Minmize $6x_1$+$5x_2$

Subject to x_1-$x_2 \leq 3$;

x_1+$2x_2 \leq 6$;

$x_1 \geq 0$;

$x_2 \geq 0$.



The smaller the point of intersection, the smaller the value that $6x_1 + 5x_2$ takes.

Graphical Method to Address Linear Optimization

Question: If we replace maximum by minimize, then what the red line is?

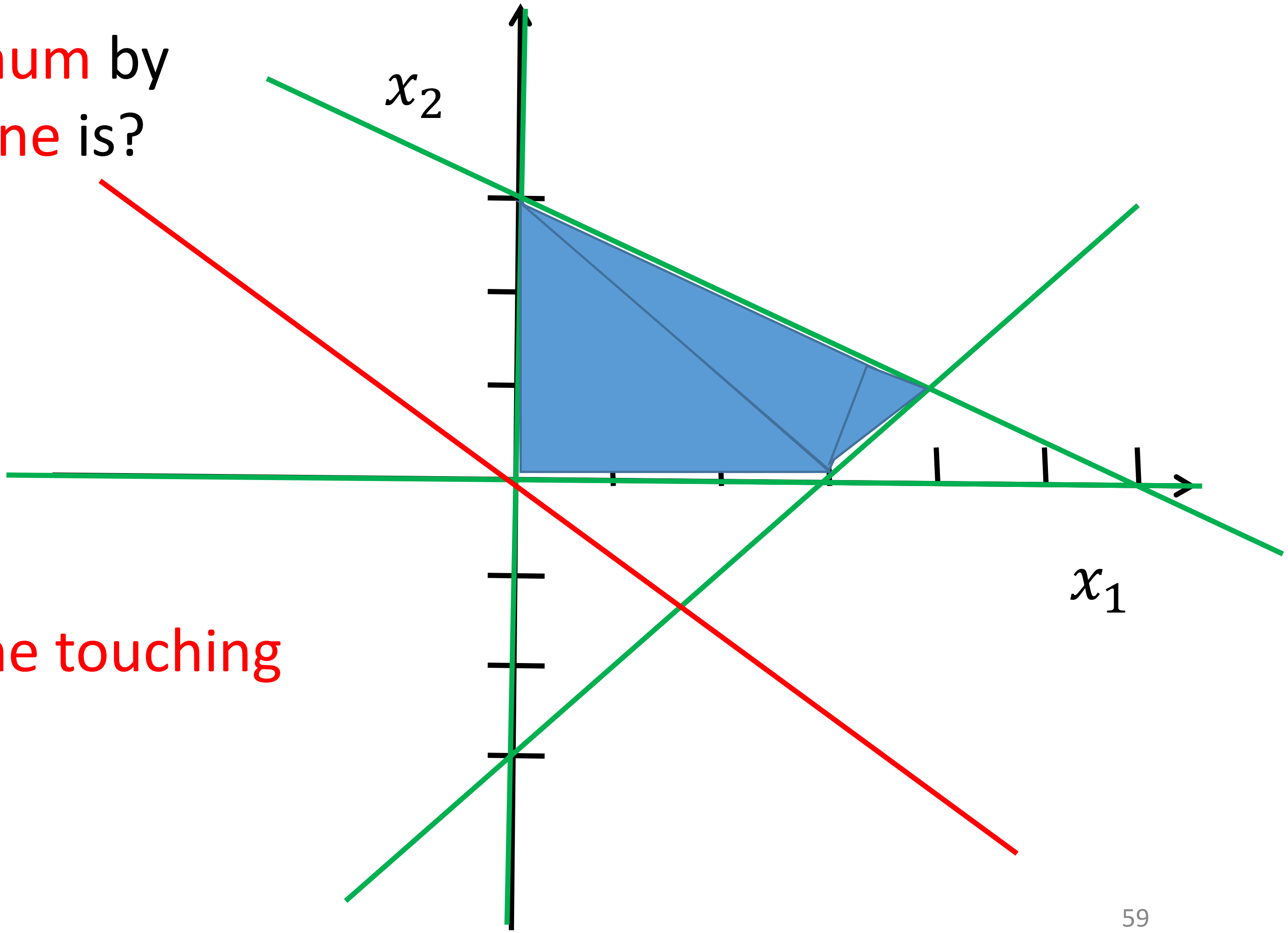
Minmize $6x_1$+$5x_2$

Subject to x_1-$x_2 \leq 3$;

x_1+$2x_2 \leq 6$;

$x_1 \geq 0$;

$x_2 \geq 0$.



So we should select the red line touching the point (0,0).

Graphical Method to Address Linear Optimization

Step 5: Finally, find out the optimum point.

An optimum point always lies on one of the corners of the feasible region. How to find it? Place a ruler on the graph sheet, parallel to the objective function. Be sure to keep the orientation of this ruler fixed in space. We only need the direction of the straight line of the objective function. Now begin from the far corner of the graph and tend to slide it towards the origin.

Graphical Method to Address Linear Optimization

Step 5: Finally, find out the optimum point.

- If the goal is to minimize the objective function, find the point of contact of the ruler with the feasible region, which is the closest to the origin. This is the optimum point for minimizing the function.

- If the goal is to maximize the objective function, find the point of contact of the ruler with the feasible region, which is the farthest from the origin. This is the optimum point for maximizing the function.

How to understand above sentences?
Please see the examples in next two pages.

Credits from topper.com

Graphical Method to Address Linear Optimization

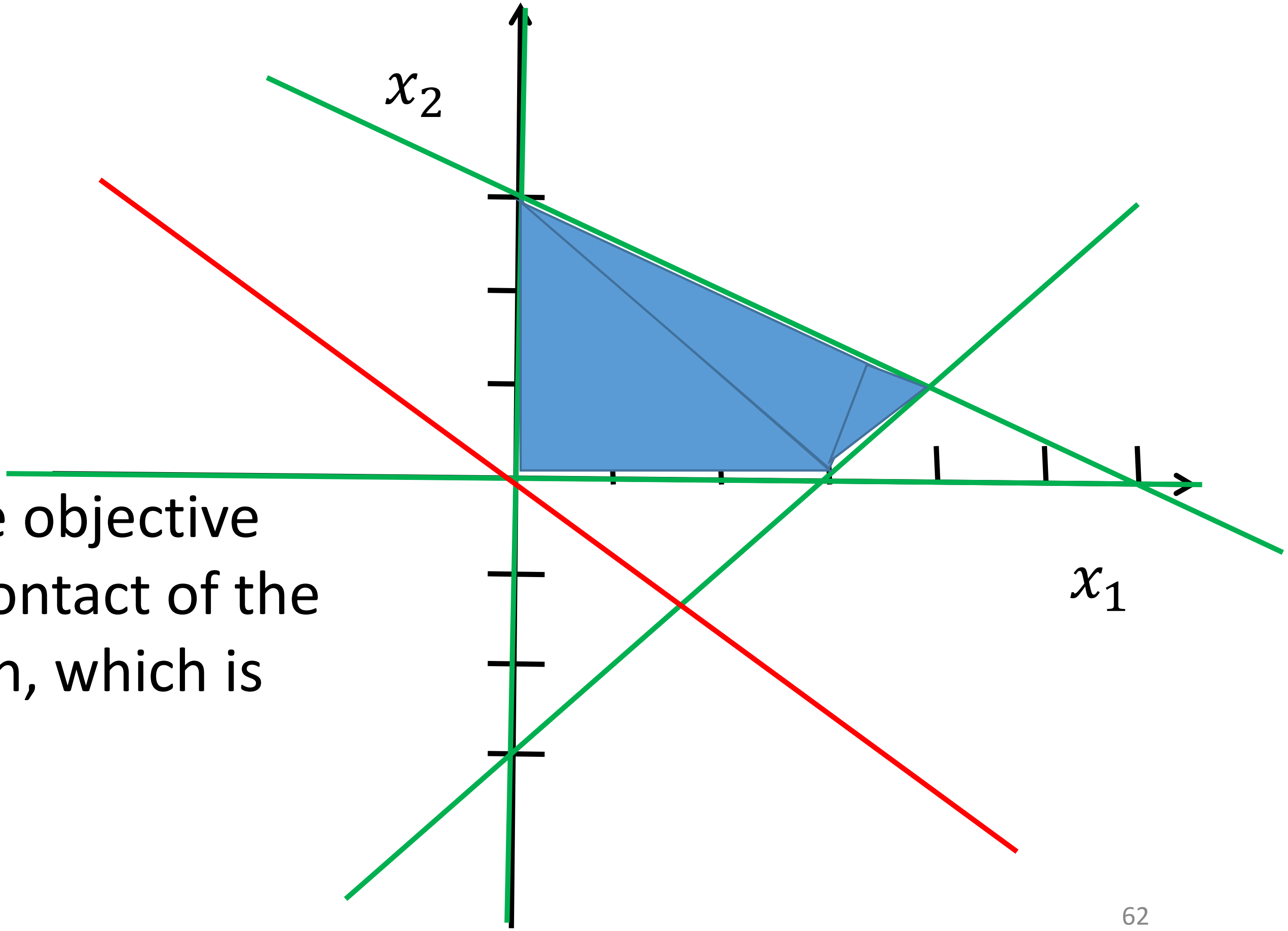
Minmize $6x_1$+$5x_2$

Subject to x_1-$x_2 \leq 3$;

x_1+$2x_2 \leq 6$;

$x_1 \geq 0$;

$x_2 \geq 0$.



- If the goal is to minimize the objective function, find the point of contact of the ruler with the feasible region, which is the closest to the origin.
- Optimum point is (0,0)

Graphical Method to Address Linear Optimization

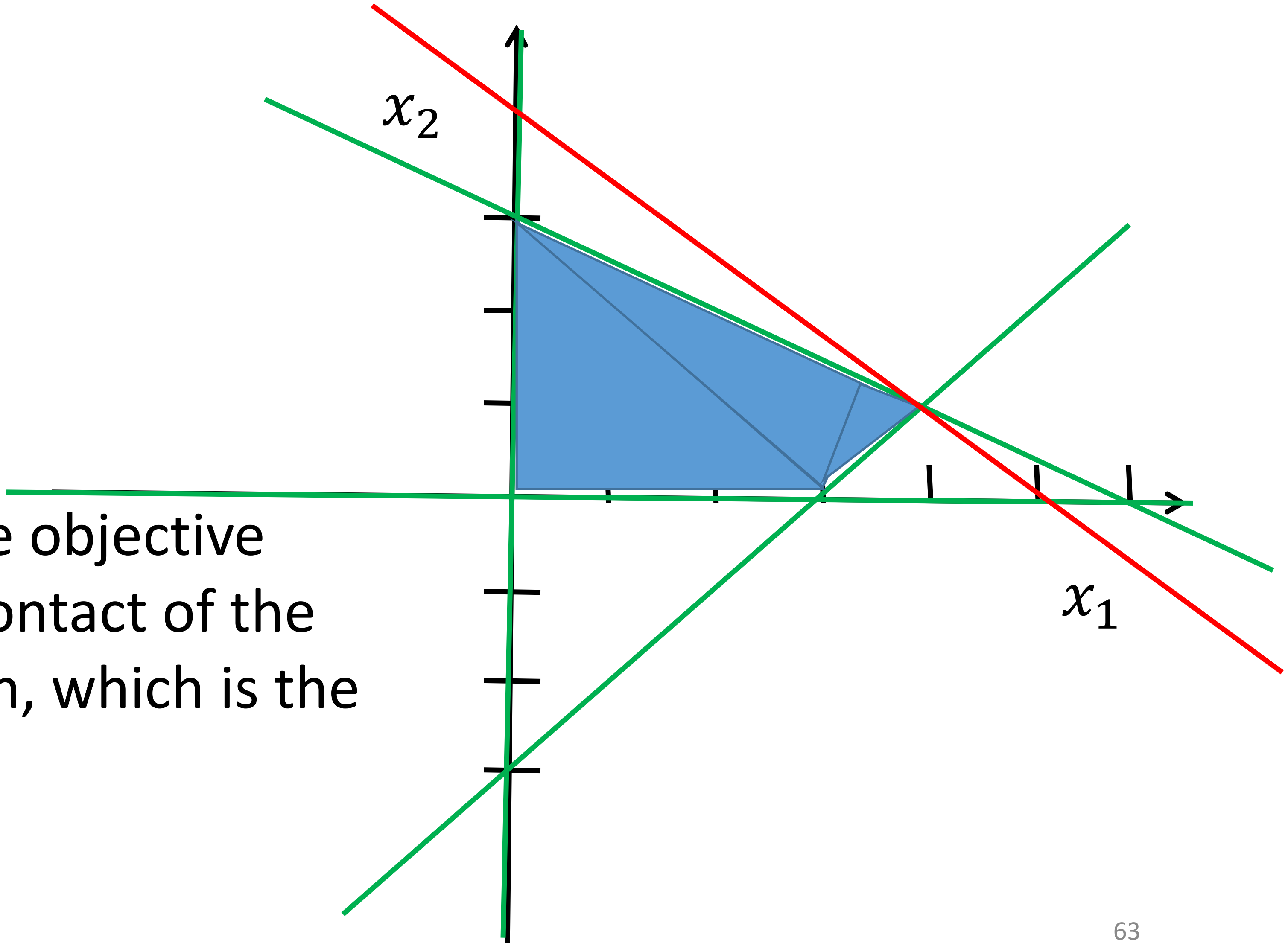
Maximum $6x_1+5x_2$

Subject to $x_1-x_2 \leq 3;$

$x_1+2x_2 \leq 6;$

$x_1 \geq 0;$

$x_2 \geq 0.$



- If the goal is to maximize the objective function, find the point of contact of the ruler with the feasible region, which is the farthest from the origin.
- Optimum point is (4,1)

Graphical Method: Exercise

A health-conscious (健康意识) family wants to have a very well controlled vitamin C-rich mixed fruit-breakfast which is a good source of dietary fibre as well; in the form of 5 fruit servings per day. They choose apples and bananas as their target fruits, which can be purchased from an online vendor (小贩) in bulk at a reasonable price.

Bananas cost 30 rupees per dozen (6 servings) and apples cost 80 rupees per kg (8 servings). Given: 1 banana contains 8.8 mg of Vitamin C and 100-125 g of apples i.e. 1 serving contains 5.2 mg of Vitamin C.

Every person of the family would like to have at least 20 mg of Vitamin C daily but would like to keep the intake under 60 mg. How much fruit servings would the family have to consume on a daily basis per person to minimize their cost?

Credits from topper.com

Graphical Method: Exercise

Solution : We begin step-wise with the formulation of the problem first.

The constraint variables – ‘x’ = number of banana servings taken and ‘y’ = number of servings of apples taken. Let us find out the objective function now.

- Cost of a banana serving = 30/6 rupees = 5 rupees. Thus, the cost of ‘x’ banana servings = 5x rupees

- Cost of an apple serving = 80/8 rupees = 10 rupees. Thus the cost of ‘y’ apple servings = 10y rupees

- Total Cost C = 5x + 10y

Credits from topper.com

Graphical Method: Exercise

Constraints: $x \geq 0; y \geq 0$ (non-negative number of servings)

Total Vitamin C intake:

$8.8x + 5.2y \geq 20 \quad (1)$

$8.8x + 5.2y \leq 60 \quad (2)$

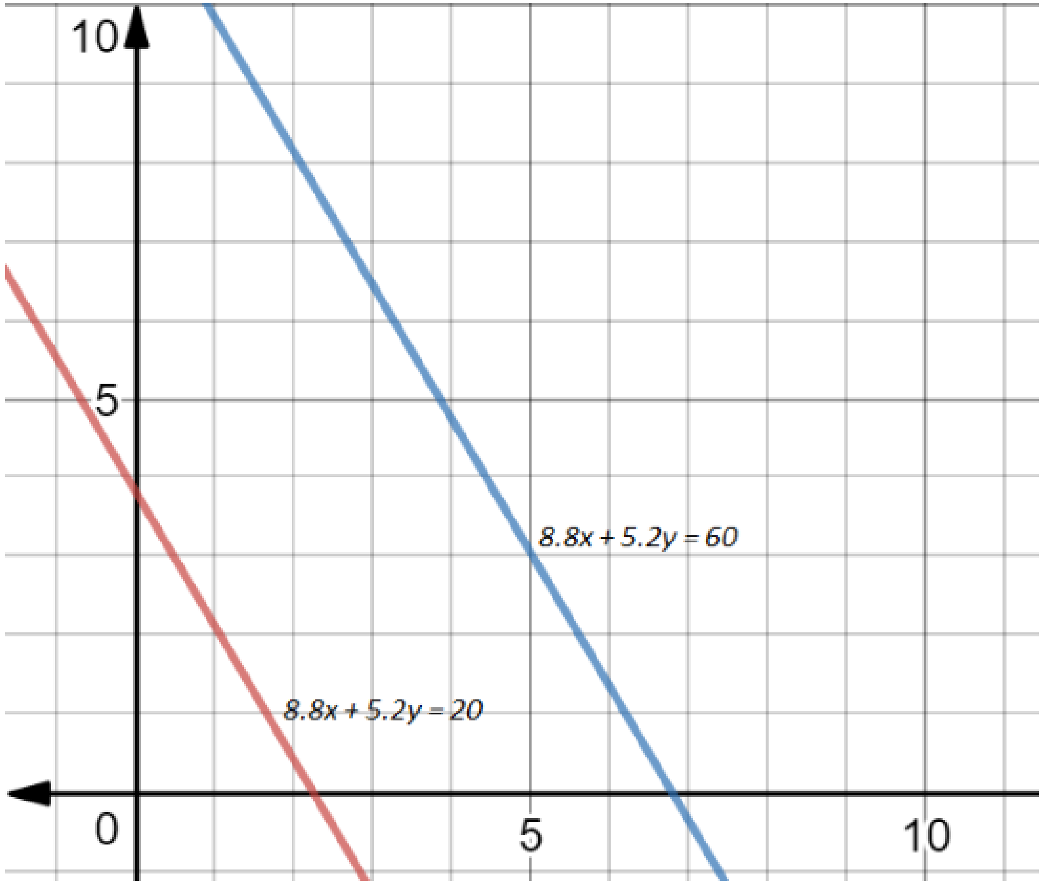
So the optimize problem is

$$\begin{aligned} &\text{Minimize } 5x + 10y \\ &\text{Subject to } 8.8x + 5.2y \geq 20 \\ &\qquad 8.8x + 5.2y \leq 60 \\ &\qquad x \geq 0; y \geq 0 \end{aligned}$$

Credits from topper.com

Graphical Method: Exercise

Now let us plot a graph with the constraint equations

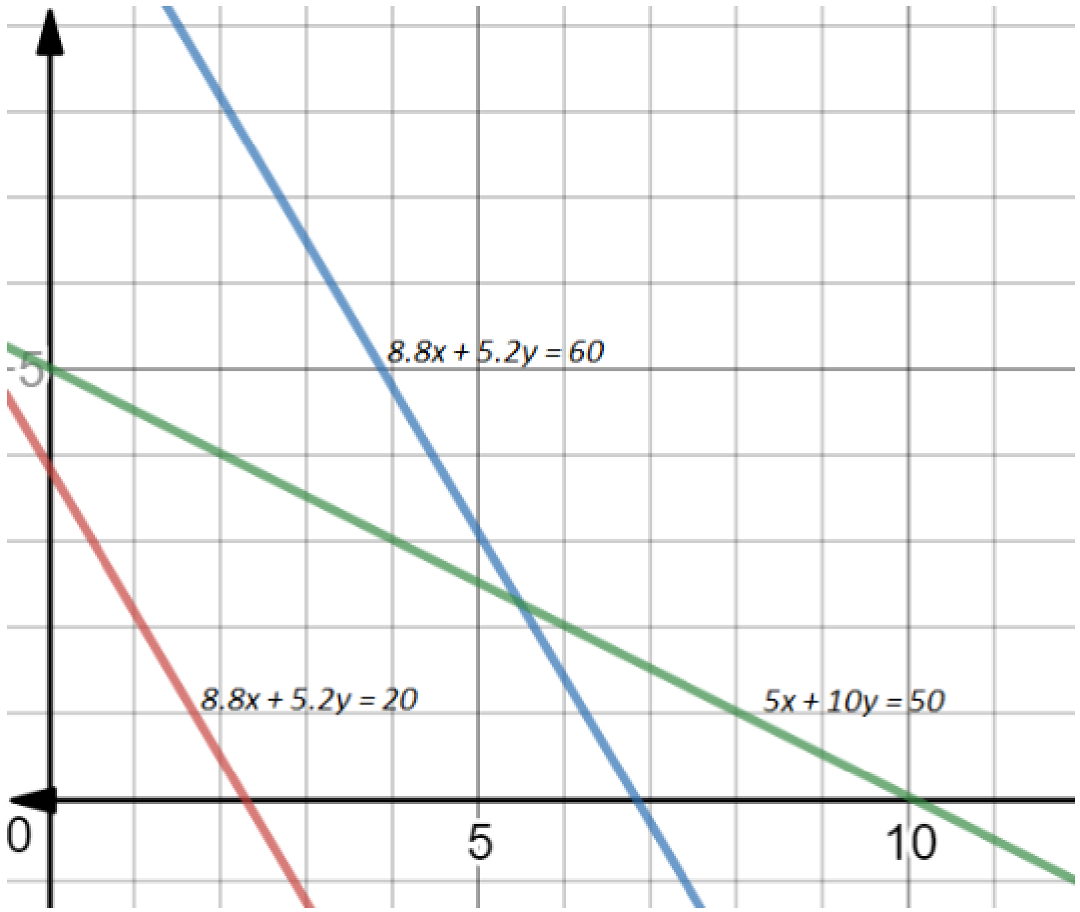


Credits from topper.com

Graphical Method: Exercise

To check for the validity of the equations, put x=0, y=0 in (1). Clearly, it doesn't satisfy the inequality. Therefore, we must choose the side opposite to the origin as our valid region. Similarly, the side towards origin is the valid region for equation 2)

Feasible Region: As per the analysis above, the feasible region for this problem would be the one in between the red and blue lines in the graph! For the direction of the objective function; let us plot 5x+10y = 50.



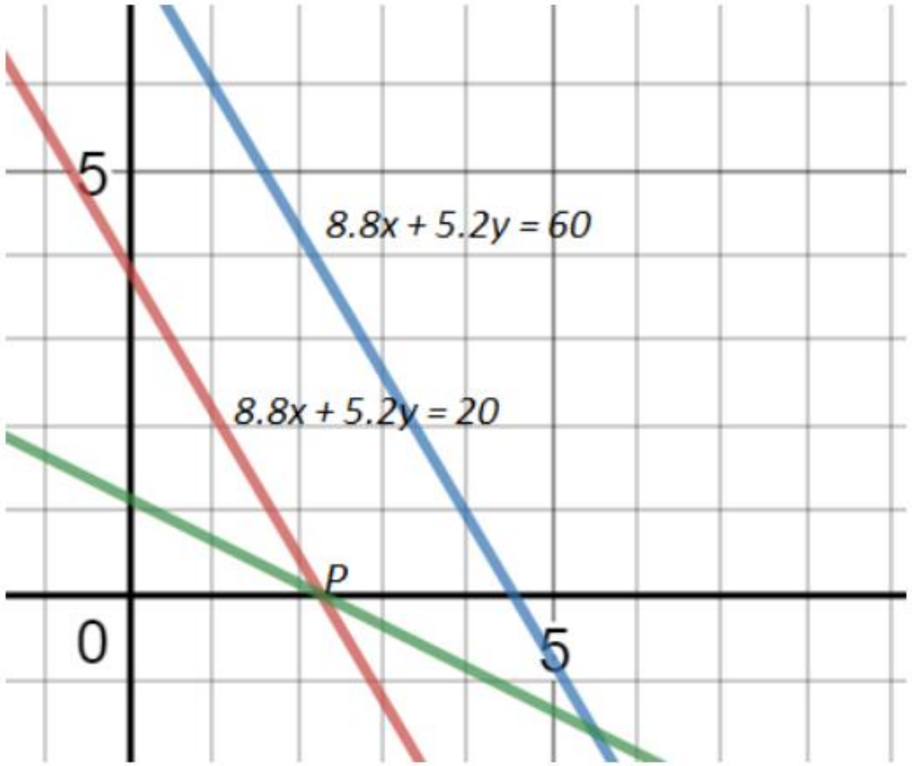
Credits from topper.com

Graphical Method: Exercise

Now take a ruler and place it on the straight line of the objective function. Start sliding it from the left end of the graph. What do we want here? We want the minimum value of the cost i.e. the minimum value of the optimum function C. Thus we should slide the ruler in such a way that a point is reached, which:

1) lies in the feasible region
2) is closer to the origin as compared to the other points

This would be our optimum point. I've marked it as P in the graph. It is the one which you will get at the extreme right side of the feasible region here. I've also shown the position in which your ruler needs to be to get this point by the line in green.



Credits from topper.com

Graphical Method: Exercise

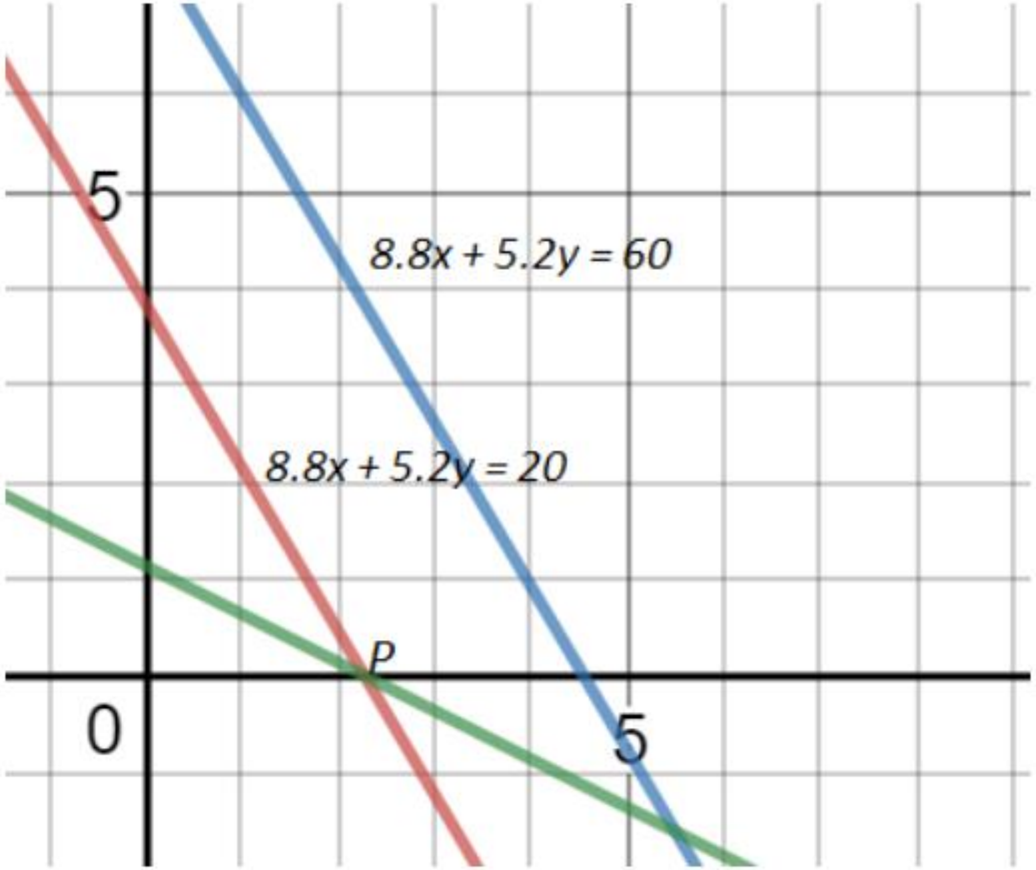
Calculating the coordinates of optimum point.

To do this, just solve the simultaneous pair of linear equations:

$$y = 0$$
$$8.8x + 5.2y = 20$$

We'll get the coordinates of 'P' as (2.27, 0).
So the optimum point is (2.27,0).

This implies that the family must consume 2.27 bananas and 0 apples to minimize their cost and function according to their diet plan.

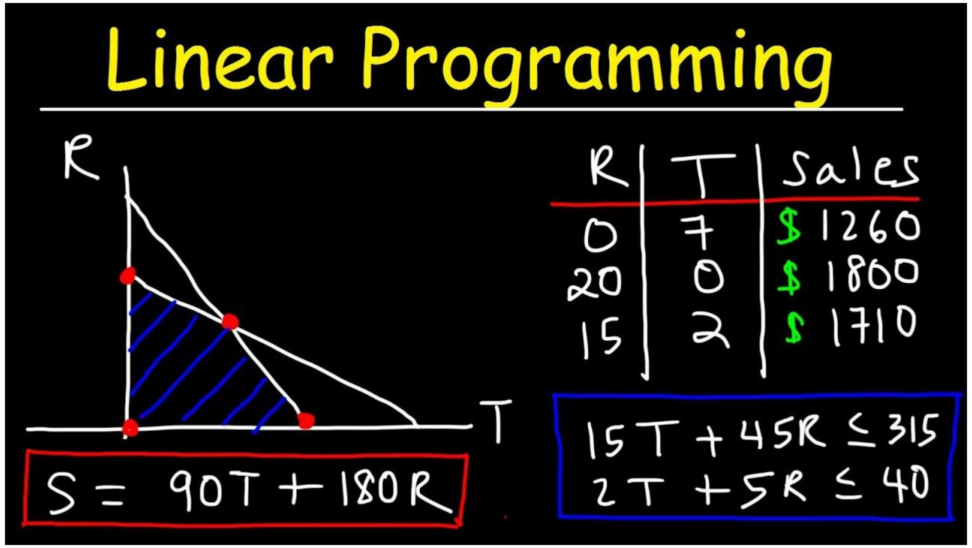


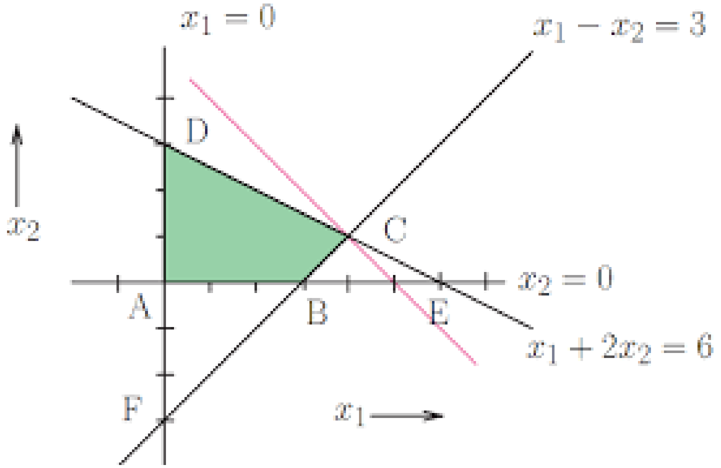
Credits from topper.com

Graphical Method: Summary

What is the purpose of a graphical method?

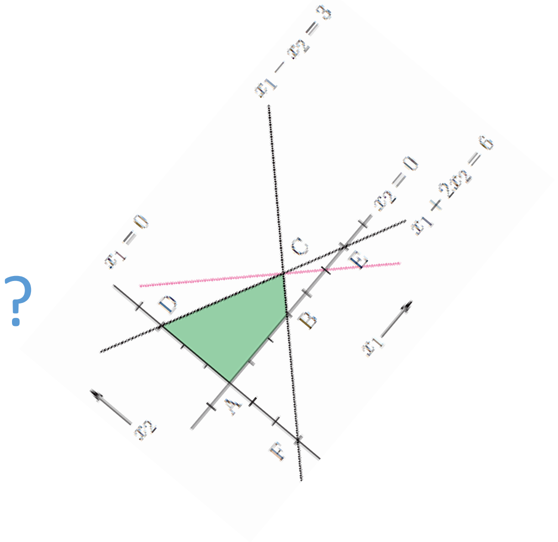
Answer: We use a graphical method of linear programming for solving the problems by finding out the maximum or minimize point of the intersection on a graph between the objective function line and the feasible region.





Credits from topper.com

Graphical Method: Summary



How do you solve the LP with the help of a graphical method?

1st Step: First of all, formulate the LP problem.

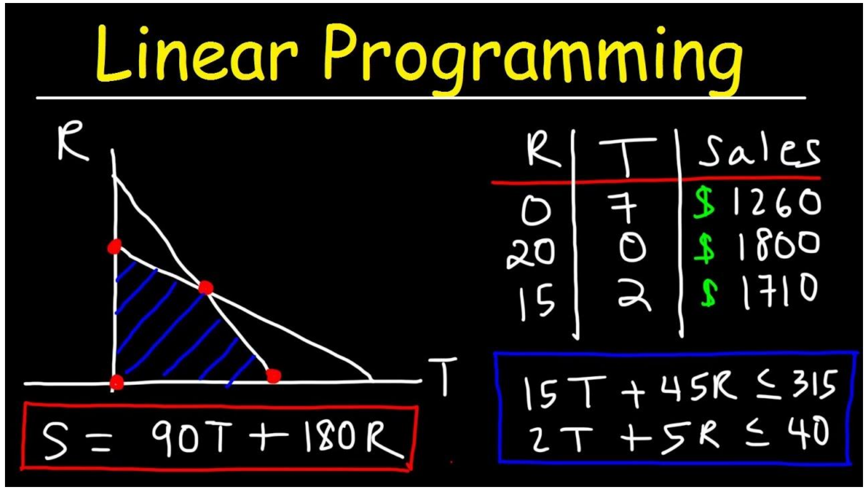
2nd Step: Then, make a graph and plot the constraint lines over there.

3rd Step: Determine the valid part of each constraint line.

4th Step: Recognize the possible solution area.

5thStep: Place the objective function in the graph.

6th Step: Finally, find out the optimum point.



Credits from topper.com

Other Methods to Address LP Problem

- Simplex Method (单纯形法） Roughly speaking, the idea of the simplex method is to represent an LP problem as a system of linear equations, and then a certain solution of the obtained system would be an optimal solution of the initial LP problem (if any exists). The simplex method defines an efficient algorithm of finding this specific solution of the system of linear equations.

https://en.wikipedia.org/wiki/Simplex_algorithm

https://www.youtube.com/watch?v=K7TL5NMlKIk

Nonlinear Optimization

- Nonlinear optimization problem: the objective $f(x)$ is nonlinear, or the any of the inequality constraints $g_i(x) \leq b_i$, i = 1, 2, …, c, or equality constraints $h_j(x) = a_j$, j = 1, 2, …, d, are nonlinear functions of the vector of variables x.

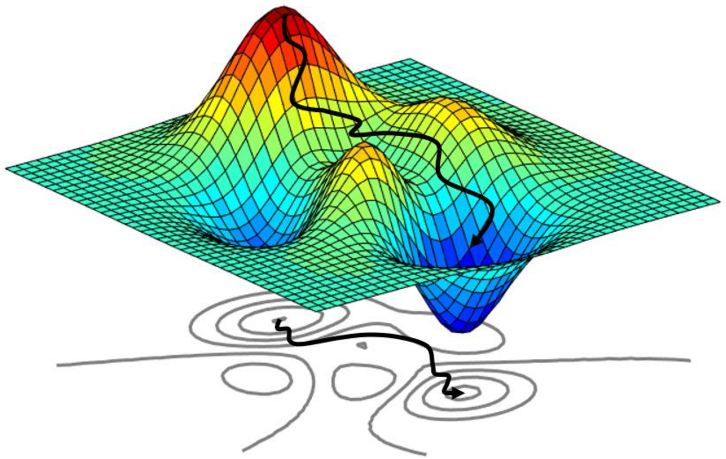
Minimize $f(x_1, x_2, \ldots, x_n)$
Subject to

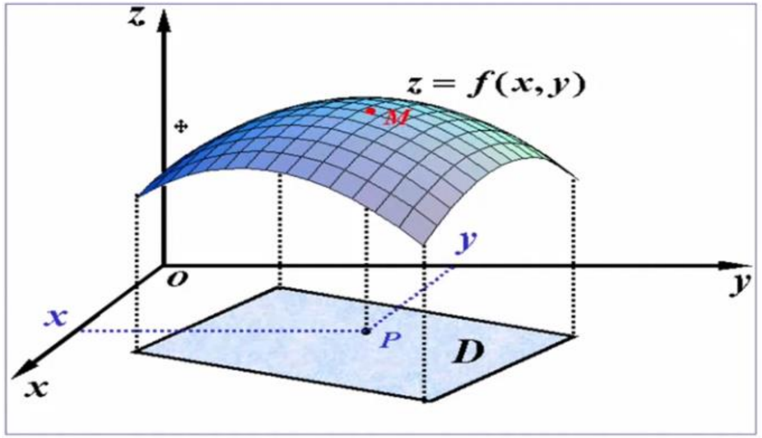
$$h_j(x_1, \ldots, x_n) = a_j, \; j = 1, \ldots, d$$

$$g_i(x_1, \ldots, x_n) \leq b_i \;, \; i = 1, \ldots., c$$

Nonlinear Optimization

- In short, nonlinear optimization is the process of solving an optimization problem where some of the constraints or the objective function are nonlinear.

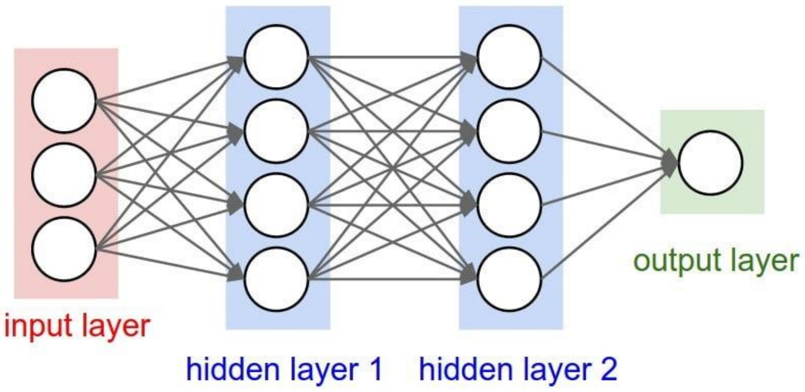


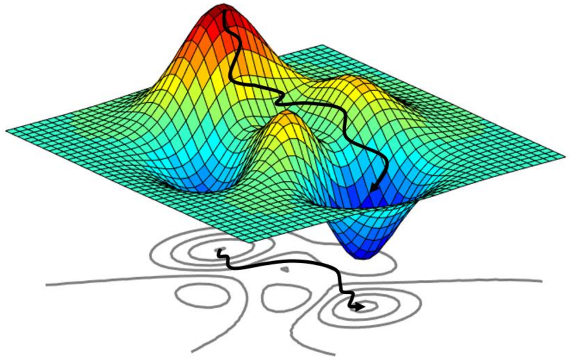


Nonlinear Optimization: Examples

- Deep neural networks with nonlinear activation function.

If fixing the parameters (non-trival parameters), deep neural network with nonlinear activation function form a nonlinear function, which implies that the objective function with respact to this deep neural network is a nonlinear function.





Nonlinear Optimization: Examples

- Locate airport while minimizing average distance.

Locate N airports each within a specified distance of a city centre, and minimise the sum of square of the distances between all the airports.

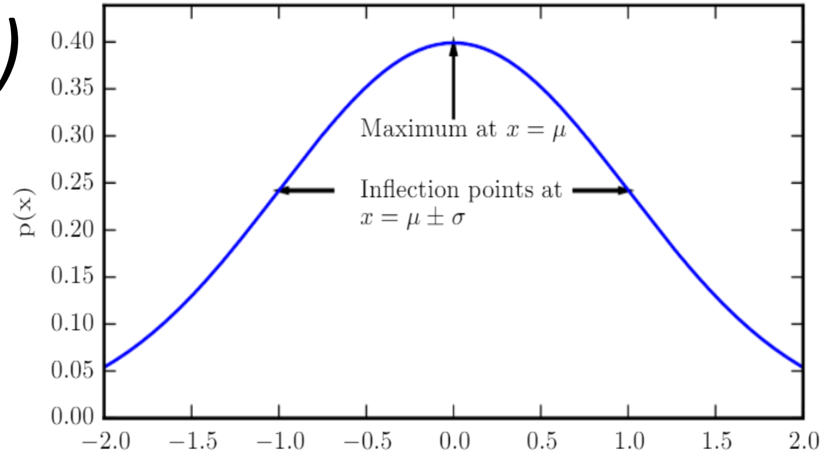
Because the sum of square of the distances is not linear function

Nonlinear Optimization: Examples

Reconsider the Maximum likelihood estimation

Suppose you have x1,x2,...,xn (i.i.d) $N(\mu,\sigma^2)$

$$\sqrt{\frac{1}{2\pi\sigma^2}}\exp(-\frac{1}{2\sigma^2}(\mathrm{x}-\mu)^2)$$



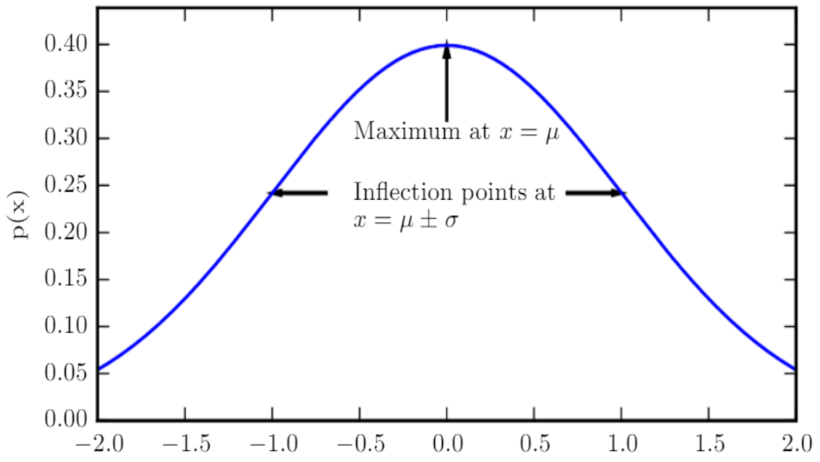
- Assume that you know σ^2
- But you don't know μ

MLE: For which μ is x1, x2, ..., xn most likely?

Nonlinear Optimization: Examples

Compute the MLE $\underset{\mu \in R}{\operatorname{argmax}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \log p_X(\mathrm{xi}; \mu)$

$$\underset{\mu \in R}{\arg\max} \frac{1}{\sqrt{2\pi}\ \sigma} \sum_{i=1}^{n} - \frac{(\mathrm{xi} - \mu)^2}{2\sigma^2}$$



Above optimization is a nonlinear optimization problem

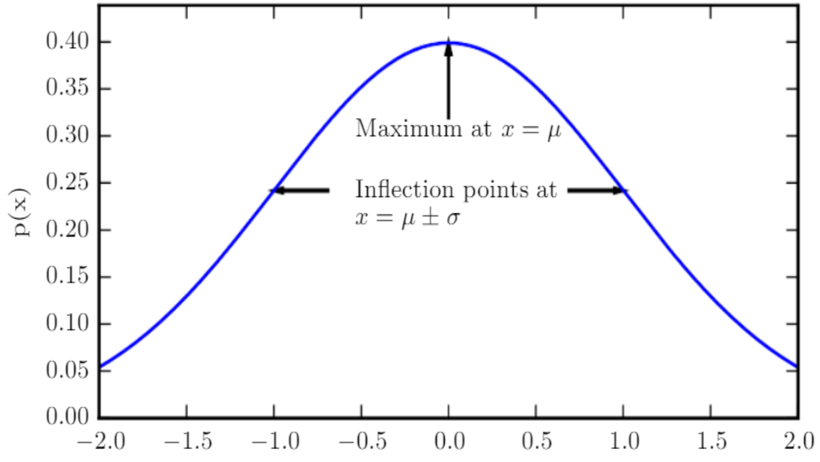
Nonlinear Optimization: Examples

- Reconsider the Maximum A Posteriori (MAP) Estimation

Suppose you have x1,x2,...,xn (i.i.d) $N(\mu,\sigma^2)$ with density

$$\mathrm{p}(\mathrm{x}|\,\mu) = \sqrt{\frac{1}{2\pi\sigma^2}}\exp(-\frac{1}{2\sigma^2}(\mathrm{x}-\mu)^2)$$

$$\mathrm{p}(\mu) = \sqrt{\frac{1}{2\pi\sigma^2}}\exp(-\frac{1}{2\sigma^2}(\mu-\mu_0)^2)$$



MAP: For which μ is?

Credicts from Dr. Liu Yang

Nonlinear Optimization: Examples

Using maximum a poserior estimation $\underset{\mu}{\operatorname{argmax}} \prod_{\mathrm{i}=1}^{\mathrm{n}} p(\mathrm{xi}|\ \mu)\mathrm{p}(\mu)$

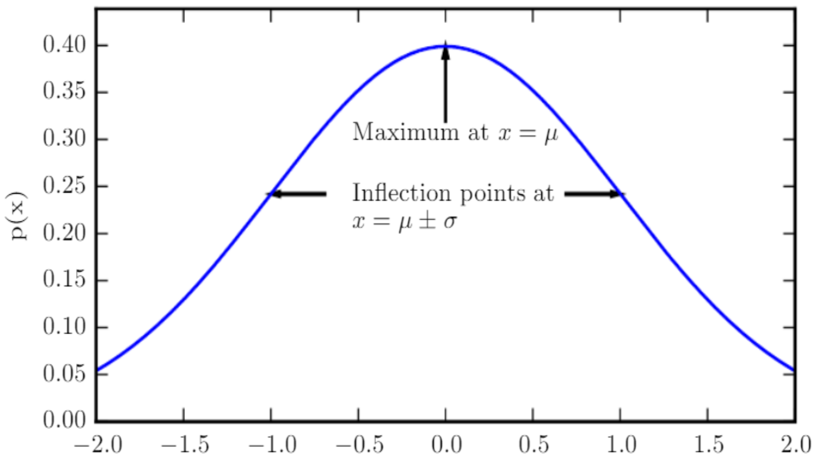
$$= \underset{\mu}{\operatorname{argmax}} \prod_{\mathrm{i}=1}^{\mathrm{n}} \exp\left(-\frac{1}{2\sigma^2}(\mathrm{xi}-\mu)^2\right)\exp\left(-\frac{1}{2\sigma^2}(\mu-\mu_0)^2\right)$$

$$= \underset{\mu}{\operatorname{argmax}} \log\left(\prod_{\mathrm{i}=1}^{\mathrm{n}} \exp\left(-\frac{1}{2\sigma^2}(\mathrm{xi}-\mu)^2\right)\exp\left(-\frac{1}{2\sigma^2}(\mu-\mu_0)^2\right)\right)$$

$$= \underset{\mu}{\operatorname{argmax}} \; -\sum_{i=1}^{n}(\mathrm{xi}-\mu)^2-(\mu-\mu_0)^2$$

$$= \underset{\mu}{\operatorname{argmin}} \; \sum_{i=1}^{n}(\mathrm{xi}-\mu)^2+(\mu-\mu_0)^2$$

- Above optimization is a nonlinear optimization problem



Nonlinear Optimization: Counter Example

Reconsider the Maximum likelihood estimation

Tossing a coin. If the possibility to appear the head is μ, then fliping a coin is a Bernoulli Distribution with parameter μ.

Bernoulli(μ)

P(X=1) = μ

P(X=0) = 1-μ

X is the random variable:

X=1 means the head appears; X=0 means the tail appears.

Nonlinear Optimization: Counter Example

Suppose that $x_1, x_2, . . ., x_n$ (i.i.d) represent the outcomes of n independent Bernoulli trials (for example, coin flipping), each with success probability μ.

- $P(X=1;\mu) = \mu$ (x=1 means the head)
- $P(X=0;\mu) = 1-\mu$ (x=0 means the tail)

So $P(X=x_i;\mu) = \mu^{x_i}(1-\mu)^{1-x_i}$

MLE: For which μ is $x_1, x_2, \ldots x_n$ most likely?

Credicts from Dr. Liu Yang

Nonlinear Optimization: Counter Example

Maximum Likelihood (ML) Estimation:

$$\underset{0\leq\mu\leq 1}{\operatorname{argmax}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \log P(X = x_{\mathrm{i}}; \mu)$$

$$\sum_{\mathrm{i}=1}^{\mathrm{n}} \log P(X = x_{\mathrm{i}}; \mu) = \sum_{\mathrm{i}=1}^{\mathrm{n}} \log \mu^{\mathrm{x_i}}(1-\mu)^{1-\mathrm{x_i}}$$

$$= \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{x_i} \log\mu + \sum_{i=1}^{n} (1-\mathrm{x_i})\log(1-\mu)$$

Above optimization is a linear optimization problem

Nonlinear Optimization: Exercise

A company wishes to sell a laptop to compete with other high-end products

- It has invested 5 million dollors to develop this product
- The success of the product will depend on the investment on the marketing campaign and the final price of the laptop
- Two important decisions:

a : amount to invest in the marketing campaign

p : price of the laptop

Nonlinear Optimization: Exercise

- Formula used by the marketing department to estimate the sales of the new product during the coming year:

$$S = 50000+5\sqrt{a}-80p$$

- The production cost of the phone is 500 dollors/unit

Whether maximizing its profits for the coming year is a nonlinear optimization problem?

Nonlinear Optimization: Exercise

Solution:

Profits from sales: $(50000+5\sqrt{a}-80p)p$

Total production costs: $(50000+5\sqrt{a}-80p)500$

Development costs: 5000000

Marketing costs: a

Total profit: $(50000+5\sqrt{a}-80p)(p-500)-5000000-a$

Nonlinear Optimization: Exercise

Solution:

maximum $(50000+5\sqrt{a}-80p)(p-500)-5000000-a$

So it is a nonlinear optimization problem.

Thank You!