Christopher Lum
lum@uw.edu

Lecture10a
Introduction to Optimization

Lecture is on YouTube

The YouTube video entitled ‘Introduction to Optimization’ that covers this lecture is located at https://youtu.be/lBXdFu6Rwn4.

Outline

-A Brief Introduction to Optimization

A Brief Introduction to Optimization

Optimization is one of the most important fields of mathematics. At its most basic, mathematical optimization involves the selection of a “best” element from a set of available choices.

The reason optimization is so important is that it can be applied to practically every engineering and science discipline. In fact, almost every problem in science and engineering (and in life in general for that matter) can be framed as an optimization problem. Philosophically, this is a reflection of the human experience, aren’t we all just trying to do the best we can with what we’ve got?

Example: Dog Selecting Food
Show example of Gus trying to eat the most tasty food available to him (while unleashed and leashed) to illustrate unconstrained and constrained optimization.

Notation

Note that the textbook gives a somewhat brief coverage on optimization. Some other more comprehensive references include:

“Convex Optimization” by Stephen Boyd and Lieven Vandenberghe
“Nonlinear Programming” by Dimitri P. Bertsekas

We will use notation from these texts instead of following the Kreyszig text. Some deviations

Kreyszig	Lecture Notes	Item/Description
$\boldsymbol{x}$	x	decision vector (note that we drop the bar notation)
$\boldsymbol{X_0}$	x^*	candidate optimal location or stationary point

Note that much of the following material follows the Bertsekas text and not the Kreyszig text.

In an **optimization problem**, the goal is to optimize (maximize or minimize) some function f. Typically, if one is attempting to maximize something (the number of cars produced from a factory), the function f is called the **objective function**. Conversely, if one is trying to minimize something (the cost to produce a single car), f is typically referred to as the **cost function**.

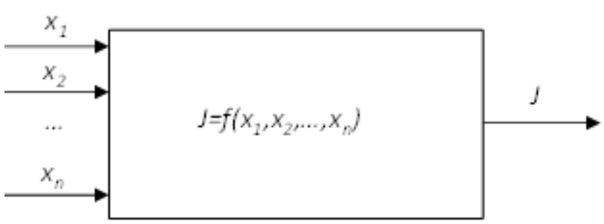
Note that many text and references use the symbol J to represent the value of the cost or objective function.

$$J = f(x)$$

In most optimization problems, the objective function or cost function f depends on several variables

$$x_1, x_2, \ldots, x_n$$

These are called **control variables** or **decision variables** because we can control them or make decisions about their values in an attempt to solve the optimization problem.



The goal in optimization is to choose $x_1, x_2, \ldots, x_n$ to maximize or minimize J

Example 1: Aircraft endurance is a function of

x_1 = fuel supply (gallons)
x_2 = drag coefficient
x_3 = air density
x_4 = total aircraft weight

Example 2: Cost to commute to work is a function of

x_1 = distance to commute
x_2 = method of transportation
x_3 = current fuel prices

Note that the example 1 is an objective/utility functions (one would typically want to maximize this function). Example 2 is typically referred to as a cost function (one would typically want to minimize this function). As we will show later, the difference between objective and cost function is simply a matter of perspective and does not influence the mathematical formulation of the problem. In other words, maximization problems can be converted to a minimization problems by simply multiplying the objective function by -1. The majority of optimization theory discusses minimization problems.

In many situations, the choice of the decision variables is not entirely free. These may be subject to some **constraints**. For instance, in example 1, all 4 decision variables cannot be negative. Furthermore, we know that the air density must be below a certain maximum value (likely the maximum air density at sea level). We will see later that these constraints will take the form of equations, inequalities, or other forms.

These restrictions define the **feasible set**. The feasible set is simply all decision variable values that respect/satisfy all the problem's constraints.

Example 3: Cost to produce a baked good is a function of

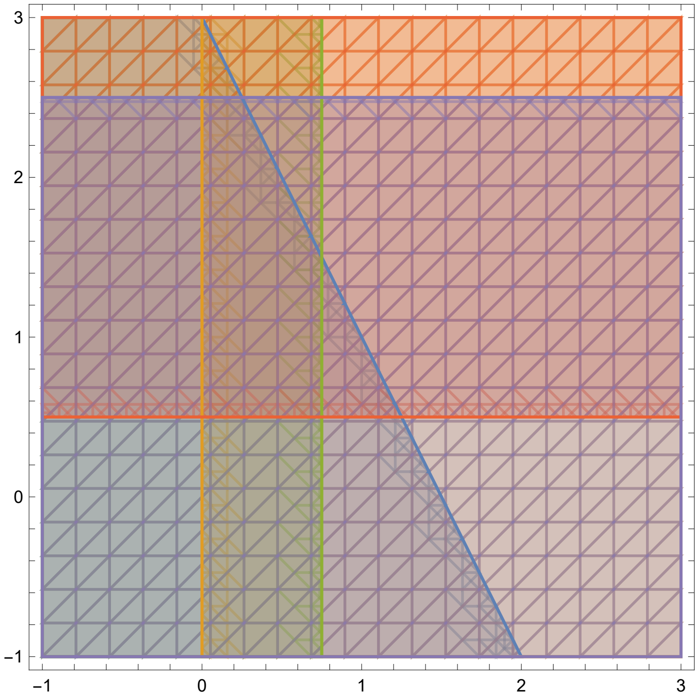
x_1 = amount of ingredient 1
x_2 = amount of ingredient 2

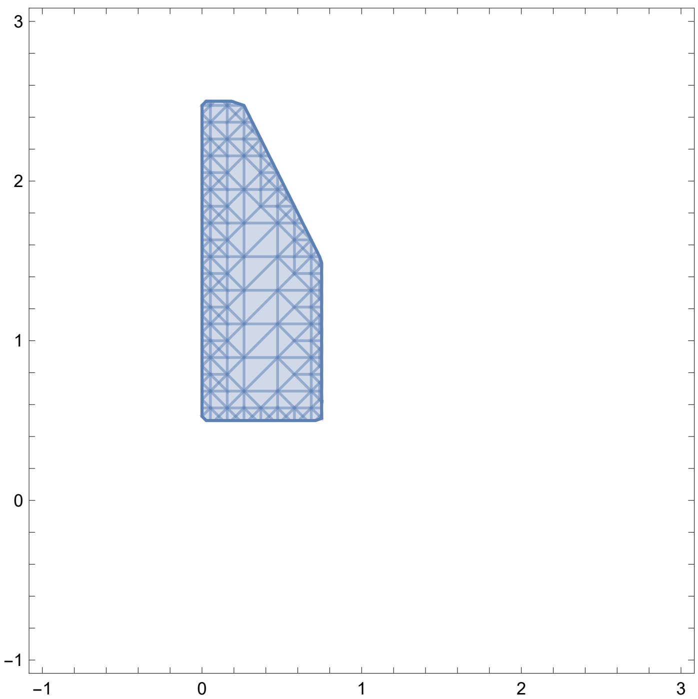
There are constraints associated with x_1 and x_2

$x_1 \geq 0$	(cannot be negative)
$x_1 \leq 0.75$	(there is only 0.75 units of ingredient 1 available)
$x_2 \geq 0.5$	(needs to have at least 0.5 units of ingredient 2 to meet FDA requirements)
$x_2 \leq 2.5$	(there is only 2.5 units of ingredient 2 available)
$2\,x_1 + x_2 < 3$	(proportion of x_1 and x_2 needs to be below a certain threshold)

```
RegionPlot[
 {2 x1 + x2 < 3,
  0 ≤ x1, x1 ≤ 0.75,
  0.5 ≤ x2, x2 ≤ 2.5},
 {x1, -1, 3}, {x2, -1, 3}]

RegionPlot[
 2 x1 + x2 < 3 &&
  0 ≤ x1 && x1 ≤ 0.75 &&
  0.5 ≤ x2 && x2 ≤ 2.5,
 {x1, -1, 3}, {x2, -1, 3}]
```





It is convenient to collect the decision variables into a vector, often called the **decision vector**. Note that most texts will drop the bar notation as it is understood that $\overline{x}$ is a vector of decision variables.

$$\overline{x} = x = \begin{pmatrix} x_1 & x_2 & \ldots & x_n \end{pmatrix} \qquad \text{(decision vector)}$$

We can define the feasible set, FS, as

$\text{FS} = \{x \mid \text{all constraints are satisfied}\}$

If there are no restrictions on the decision vector (AKA the decision/control variables), then there are no constraints on the problem and we we call this an **unconstrained optimization problem**. If there are constraints on the decision vector then this is a **constrained optimization problem**.

With this infrastructure in place, we can pose the optimization problem as

$$(\wp) \quad \text{min/max } J = f(x)$$
$$\text{such that } x \in \text{FS}$$

Example: Commute (Constrained Optimization)

Consider the following optimization problem. We would like to figure out the best way to commute from Bainbridge Island to the UW campus using an optimization framework. Consider a set of possible solutions as

z_1 = drive from house to ferry, catch ferry, drive to campus
z_2 = bus from house to ferry, catch ferry, bus to campus
z_3 = bike from house to ferry, catch ferry, bike to campus
z_4 = run to ferry, swim across Puget sound, run to campus
z_5 = rent helicopter to fly from house to campus directly

We can denote the set of all possible solutions as $Z = \{z_1, z_2, \ldots, z_5\}$ (the feasible set)

Z = feasible set

Let us now examine different cost functions and how choosing a different cost function will influence the solution to the optimization problem.

<u>Situation 1: Minimize Time</u>

Suppose that we would like to minimize the amount of time required for the commute. Therefore, the cost function would be

$J_1 = f_1(z)$ = time required for action z

Therefore, the optimization problem can be formulated as

$(\wp_1)\ \underset{z \in Z}{\text{minimize}}\ f_1(z)$ (minimize cost function $J_1(z)$ subject to z must be in the feasible set Z)

In this situation, the solution to this problem, $(\wp_1)$, would be

z_5 =rent helicopter to fly from house to campus directly

This would most likely get you to campus the fastest. However, this would cost a significant amount of money.

Situation 2: Minimize Money

The problem with cost function $J_1(z)$ is that we did not include any monetary aspects into the cost function. Therefore, another potential cost function might be

$$J_2 = f_2(z) \text{ money required for action } z$$

Therefore, the optimization problem can be formulated as

$$(\wp_2) \underset{z \in Z}{\text{minimize}}\, f_2(z)$$

In this situation, the solution to this problem, $(\wp_2)$, would most likely be

z_4 =walk to ferry, swim across Puget sound, walk to campus

This would most likely get you to campus the and spend the least amount of money. However, this would result in an unrealistic time to commute (and other factors such as likely death swimming 12 miles across Puget sound).

Situation 3: Minimize Time/Money Trade Off

So a 3rd potential cost function might be a trade off between time and money. Therefore, another potential cost function might be

$$J_3 = f_3(z) = q\, f_1(z) + r\, f_2(z)$$

where $f_1(z)$ = time required for action z
$f_2(z)$ = money required for action z
q, r = scalar coefficients to trade off between time and money

So we see that we can tune the optimization problem by choosing appropriate q and r parameters. For example

large q, small $r \Rightarrow$ problem reverts to a minimize time problem
small q, large $r \Rightarrow$ problem reverts to a minimize money problem
medium q, medium $r \Rightarrow$ trade off between time and money compromise

Therefore, the optimization problem can be formulated as

$$(\wp_3) \underset{z \in Z}{\text{minimize}}\, f_3(z)$$

In this situation, the solution to this problem, $(\wp_2)$, would most likely be

z_1 = drive from house to ferry, catch ferry, drive to campus (if time costs relatively more than money)

z_2 = bus from house to ferry, catch ferry, bus to campus (if money costs relatively more than time)

Situation 4: Minimize Time/Money Trade Off with Additional Constraints

We can add additional constraints to the problem to further tailor the behavior of the solution. For example, let us continue with our 3rd cost function (trade off between time and money) but add an additional constraint to the problem. For example, we could require that during the commute, we would like to get at least 30 minutes of exercise as well.

constraint 1 = commute yields at least 30 minutes of exercise

Therefore, the optimization problem can be formulated as

$$(\wp_4) \underset{z \in Z}{\text{minimize}} f_3(z)$$

such that $g_1(z) \geq 30$

where $g_1(z)$ = number of minutes exercised during action z

This additional constraint has the effect of making the feasible set smaller.

$g_1(z_1) \ngeq 30 \Rightarrow z_1$ is not allowed as a solution (drive from house to ferry, catch ferry, drive to campus)

$g_1(z_2) \ngeq 30 \Rightarrow z_2$ is not allowed as a solution (bus from house to ferry, catch ferry, bus to campus)

$g_1(z_3) \geq 30 \Rightarrow z_3$ is a potential valid solution (bike from house to ferry, catch ferry, bike to campus)

$g_1(z_4) \geq 30 \Rightarrow z_4$ is a potential valid solution (run to ferry, swim across Puget sound, run to campus)

$g_1(z_5) \ngeq 30 \Rightarrow z_5$ is not allowed as a solution (rent helicopter to fly from house to campus directly)

Therefore, the feasible set for $(\wp_4)$ is

$$\tilde{Z} = \{z_3, z_4\}$$

Then problem $(\wp_4)$ could be alternatively expressed as

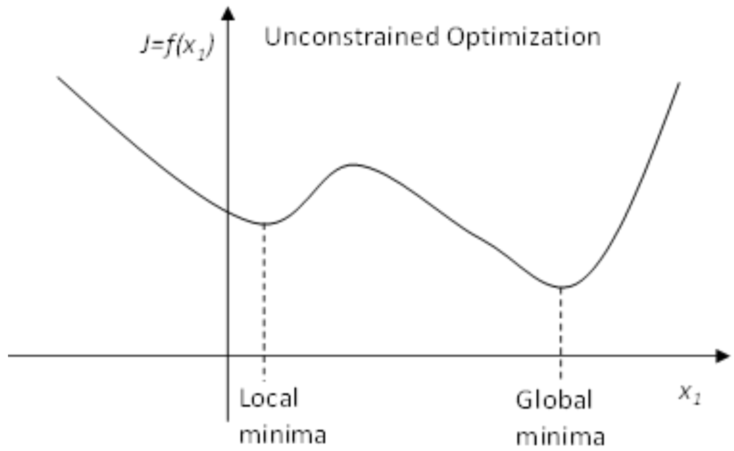
$$(\wp_4) \underset{z = \tilde{Z}}{\text{minimize}} f_3(z)$$

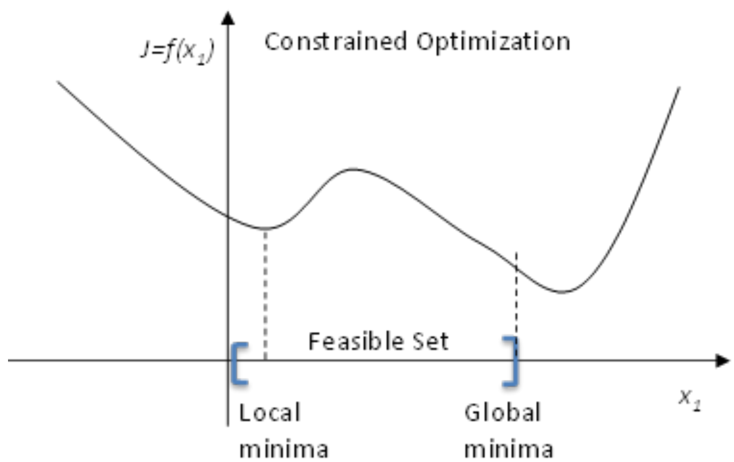
In this situation, the solution to this problem, $(\wp_4)$, would most likely be

z_3 = bike from house to ferry, catch ferry, bike to campus) (satisfies exercise constraint, good trade off between time/money)

General Picture

So in general the picture that goes along with optimization is similar to those shown below.





Future Topics

Next topics regarding optimization

-Unconstrained Optimization
-The Taylor Series
-Algorithms for Unconstrained Optimization
-Constrained Optimization
-Convex Optimization
-Applying Optimization to Engineering Problems (such as trimming an aircraft)

Entire playlist of optimization videos is located at https://www.youtube.com/playlist?list=PLxdnSsBqCrrHo2EYb_sMctU959D-iPybT

Backup

Example: Choosing an Optimal Location

Consider the scenario where a UAV is tasked with searching a region. It needs to choose a position to search

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

where x_1 = position east (x coordinate)
x_2 = position north (y coordinate)

The cost associated with a location is a combination of the distance away and how likely the target is to be located in a given cell of the map.

$$f(x) = \alpha\left(x_1^{\,2} + x_2^{\,2}\right)^{1/2} + \beta f_{\text{score}}(x_1,\, x_2) \qquad \textbf{\textit{(Eq.E.1)}}$$

where $f_{\text{score}}(x_1, x_2)$ =score of the cell at location x_1, x_2
α = relative weight/importance of distance
β = relative weight/importance of score

The UAV might only have a certain amount of fuel and therefore not all locations are feasible. We can impose a constraint stating that the UAV should only choose locations with are within a specific radius, R, of its current location. This constraint can be mathematically stated as

$$(x_1 - \text{UAV}_x)^2 + \left(x_2 - \text{UAV}_y\right)^2 \le R^2 \qquad \textbf{\textit{(Eq.E.2)}}$$

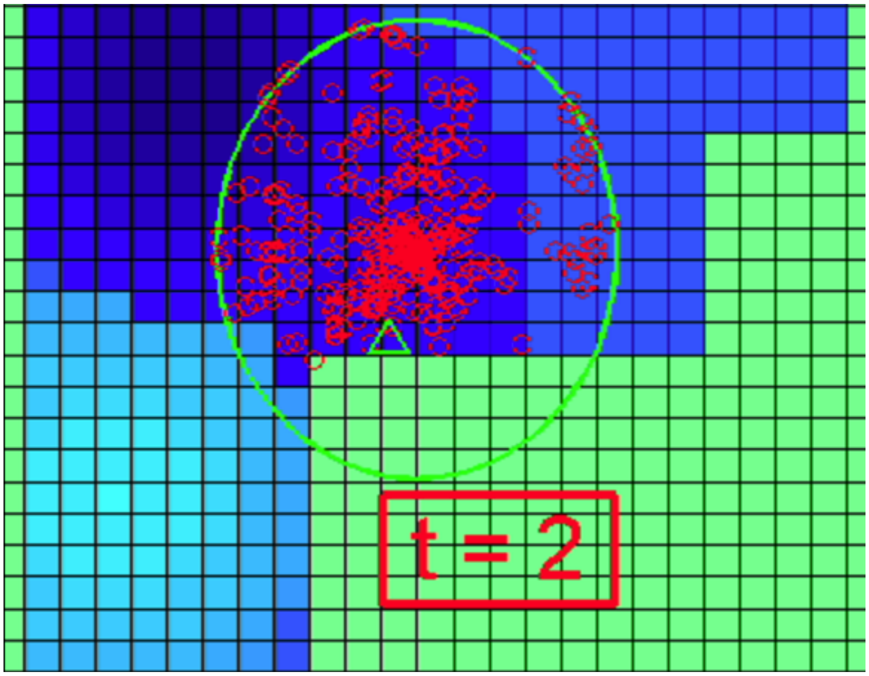
where UAV_x = current x position of the UAV
UAV_y = current y position of the UAV

This can be rewritten in standard form as

$$f_1(x_1,\, x_2) \le 0$$

where $f_1(x_1, x_2) = (x_1 - \text{UAV}_x)^2 + \left(x_2 - \text{UAV}_y\right)^2 - R^2$

Graphically, the scenario can be visualized as shown below (ignore the red dots and green triangle).



<Show the probability collectives movie to show how to approximately solve this problem>

Feasible Set