Basic Optimization Concepts

CE 377K
Stephen D. Boyles
Spring 2015

1 Introduction

At the center of every policy or planning decision are choices intended to achieve one or more outcomes. The field of "optimization" is concerned with how this process can be quantitatively modeled, and, within the bounds of these quantitative models, how the best decisions can be made. In the words of the Institute for Operations Research and Management Sciences, it is "the science of better." This field is often known as *operations research*, and has close ties with industrial or systems engineering. Although this field has its roots in the work of Fermat, Newton, and Gauss in the development of the calculus, and in the work of Joseph-Louis Lagrange in the 18th and 19th centuries, optimization truly came into its own in the twentieth century with the advent of computers which can solve large-scale optimization problems often involving tens of thousands or even millions of variables.

This field is closely related to the transportation planning world, both in an obvious and a less-obvious sense. Certain logistics and planning problems very clearly lend themselves to the "three-part" sructure of optimization problems, described in the following section.

The aim of these notes, and of this portion in the course, is to familiarize yourslef with these three parts constituting optimization problems, and to show how practical optimization problems can be translated into mathematical language and notation. As a necessary part of doing this, we will have to make decisions about what we can safely assume about the problem at hand. Pay careful attention to this: there are far too many people who jump straight into complicated models and solution methods without taking the time to properly think about what it is they are trying to solve, what kinds of assumptions are reasonable, and so forth.

2 What is an optimization problem?

Every optimization problem has three components: an objective function, decision variables, and constraints. When one talks about *formulating* an optimization problem, it means translating a "real-world" problem into the mathematical equations and variables which comprise these three components.

The objective function, often[1] denoted f or z, reflects a single quantity to be either maximized or minimized. Examples in the transportation world include "minimize congestion", "maximize safety", "maximize accessibility", "minimize cost", "maximize pavement quality", "minimize emissions," "maximize revenue," and so forth. You may object to the use of a *single* objective function,

[1] But not always! Often problem-specific notation will be used, such as c for cost, and so forth. The notation given in this section is what is traditionally used if we are referring to a generic optimization problem outside of a specific context.

since real-world problems typically involve many different and conflicting objectives from different stakeholders, and this objection is certainly valid. There are several reasons why I'm restricting the scope of these notes to single objectives. From a historical perspective, if we don't know how to optimize a single objective function, then we have no hope of being able to optimize multiple objectives simultaneously, since the latter build on the former. From a pedagogical perspective, the methods of multi-objective optimization are much more complex, and the basic concepts of optimization are best learned in a simpler context first. From a mathematical perspective, the definition of "simultaneously optimize" is very tricky, since there is probably no plan which will, say, simultaneously minimize congestion and agency cost — therefore multiobjective optimization is "fuzzier," and this fuzziness can be confusing when first taught. From a curricular perspective, the types of optimization models which show up most frequently in this course actually can be formulated as single-objective problems with no loss of real-world applicability. So, don't be alarmed by this restriction to a single objective[2], but do keep it in the back of your head if you use optimization beyond this course.

The decision variables (often, but not always, denoted as the vector $\mathbf{x}$) reflect aspects of the problem that you (or the decision maker) has control over. This can include both variables you can directly choose, as well as variables which you indirectly influence by the choice of other decision variables. For example, if you are a private toll road operator trying to maximize your profit, you can directly choose the toll, so that is a decision variable. You can't directly choose how many people drive on the road — but because that's influenced by the toll you chose, the toll road volume should be a decision variable as well. However, you should avoid including extraneous decision variables. Every decision variable in your formulation should either directly influence the objective function, or influence another decision variable that affects the objective function.

Constraints represent any kind of limitation on the values that the decision variables they take. The most intuitive types of constraints are those which directly and obviously limit the choices you can make: you can't exceed a budget, you are required by law to provide a certain standard of maintenance, you are not allowed to change the toll by more than \$1 from its current value, and so forth. One very frequent mistake is to omit "obvious" constraints (e.g., the toll can't be negative). An optimization formulation must be complete and not leave out any constraint, no matter how obvious it may seem to you.[3] The second type of constraint is required to ensure consistency among the decision variables. Following the toll road example from the previous paragraph, while you can influence both the toll (directly) and the roadway volume (indirectly), the roadway volume must be consistent with the toll you chose. This relationship must be reflected by an appropriate constraint; say, a demand function $d(\tau)$ giving the demand d for the tollway in terms of the toll τ, which can be estimated in a variety of econometric ways. Constraints are the primary way these linkages between decision variables can be captured.

[2] At least not yet.

[3] Two reasons for this. First, what is obvious to you may not be obvious to someone else; second, and perhaps more importantly, any optimization problem of reasonable size will be solved by a computer, which has no common sense whatsoever, so everything must be spelled out.

3 A few examples

The discussion to this point may be a bit abstract. The best way to demonstrate these concepts is through a few tangible examples. These examples will also demonstrate, *en passant*, the stylistic conventions for writing optimization problems. In each of these examples, the objective function, decision variables, and constraints will first be described intuitively, using plain English. Only then will we introduce mathematical notation, equations, inequalities, and so forth. You should make this a habit, no matter how simple the problem may appear at first glance. As an instructor, many of the mistakes I see in formulating optimization problems happen because a student started defining notation and equations before the whole problem is fully understood, and was thereby led down the wrong path.[4] It's much better to be patient and understand the totality of the problem before translating your intuitive descriptions into mathematical ones.

This first example may be familiar to you from introductory calculus:

Example 1. *You have 60 feet of fence available, and wish to enclose the largest rectangular area possible. What dimensions should you choose for the fenced-off area?*

Solution. The objective is clear from the problem statement: you wish to maximize the area enclosed by the fence. The decision variables are not directly given in the problem. Rather, you are told that you must enclose a rectangular area. To determine a rectangle, you need to make two decisions: its length and its width. These are both decision variables you can control directly, and there are no indirect decision variables because the length and width directly determine its area. There is one obvious constraint — the perimeter of the fence cannot exceed 60 feet — and two less obvious ones: the length and width must be nonnegative. Since the length and width are independent of each other (the perimetric constraint notwithstanding), there is no need to add a "consistency constraint" linking them.

Now we can proceed to introduce mathematical notation. Starting with the decision variables, let L represent the length and W the width. Then the objective function is to maximize the area, which is equal to LW. The perimeter of a rectangle with length L and width W is $2L + 2W$, so the perimeter constraint can be expressed $2L + 2W \leq 60$. Finally, the nonnegativity constraints are $L \geq 0$ and $W \geq 0$. All of this information is concisely represented in the following way:

$$\begin{array}{rll} \max\limits_{L,W} & LW & \\ \text{s.t.} & 2L + 2W & \leq 60 \\ & L & \geq 0 \\ & W & \geq 0 \end{array}$$

Note that the decision variables are indicated underneath the word "max" (which would naturally be replaed by "min" if this were a minimization problem), and that each constraint is listed on a separate row underneath the objective function. The letters "s.t." stand for "subject to."[5] ■

[4] There is an old adage: "if your only tool is a hammer, every problem looks like a nail." If you start putting an optimization problem into a certain framework, it is very tempting to try to force the rest of the problem into that same paradigm whether it really fits or not.

[5] There is a joke about this type of problem. A business major, an engineer, and a mathematician have a competition to see who can enclose the greatest area with a fence (this time without the requirement that the area be

Example 2. (Transit frequency setting.) *You are working for a public transit agency in a city, and must decide the frequency of service on each of the bus routes. The bus routes are known and cannot change, but you can change how the city's bus fleet is allocated to each of these routes. (The more buses assigned to a route, the higher the frequency of service.) Knowing the ridership on each route, how should buses be allocated to routes to minimize the total waiting time?*

Solution. To formulate this as an optimization problem, we need to identify an objective, decision variables, and constraints. For this problem, as we do this we will be faced with other assumptions which must be made. With this and other such problems, there may be more than one way to write down an optimization problem, what matters is that we *clearly state all of the assumptions made.* Another good guiding principle is to *start with the simplest model which captures the important behavior*, which can then be refined by relaxing assumptions or replacing simple assumptions with more realistic ones.

We need some notation, so let R be the set of bus routes. We know the current ridership on each route d^r. Here we will make our first assumption: that the demand on each route is inelastic and will not change based on the service frequency. Like all assumptions, it is not entirely true but in some cases may be close enough to the truth to get useful results; if not, you should think about what you would need to replace this assumption, which is always a good exercise. We are also given the current fleet size (which we will denote N), and must choose the number of buses associated with each route (call this n^r) — thus the decision variables in this problem are the n^r values.

The objective is to minimize the total waiting time, which is the sum of the waiting time for the passenger on each route. How long must passengers wait for a bus? If we assume that travelers arrive at a uniform rate, then the average waiting time will be half of the service headway. How is the headway related to the number of buses on the route n^r? Assuming that the buses are evenly dispersed throughout the time period we are modeling, and assuming that each bus is always in use, then the headway on route r will be the time required to traverse this route (T_r) divided by the number of buses n_r assigned to this route. So, the average delay per passenger is half of the headway, or $T_r/(2n_r)$, and the total passenger delay on this route is $(d_r T_r)/(2n_r)$. This leads us to the objective function

$$D(\mathbf{n}) = \sum_{r \in R} \frac{d_r T_r}{2n_r} \tag{1}$$

in which the total delay is calculated by summing the delay associated with each route.

What constraints do we have? Surely we must run at least one bus on each route (or else we would essentially be cancelling a route), and in reality as a matter of policy there may be some lower limit on the number of buses assigned to each route; for route r, call this lower bound L_r. Likewise, there is some upper bound U^r on the number of buses assigned to each route as well. So, we can introduce the constraint $L_r \le n_r \le U_r$ for each route r.

rectangular). The business major boldly implements his strategic vision by building a triangle. The engineer thinks for a little bit, and calculates that a circular fence actually encloses the most area. The mathematician thinks some more, builds a tiny fence, climbs inside, and says "I define myself to be on the outside!"

Putting all of these together, we have the optimization problem

$$\begin{aligned} \min_{\mathbf{n}} \quad & D(\mathbf{n}) = \sum_{r \in R} \frac{d_r T_r}{2 n_r} & \\ \text{s.t.} \quad & n_r \geq L_r & \forall r \in R \\ & n_r \leq U_r & \forall r \in R \end{aligned}$$

■

Example 3. (Scheduling maintenance.) *You are responsible for scheduling routine maintenance on a set of transportation facilities (such as pavement sections or bridges.) The state of these facilities can be described by a condition index which ranges from 0 to 100. Each facility deteriorates at a known, constant rate (which may differ between facilities). If you perform maintenance during a given year, its condition will improve by a known amount. Given an annual budget for your agency, when and where you should perform maintenance to maximize the average condition of these facilities? You have a 10 year planning horizon.*

Solution. In contrast to the previous example, where the three components of the optimization problem were described independently, from here on problems will be formulated in a more organic way, describing a model built from the ground up. (This is how optimization models are usually described in practice.) After describing the model in this way, we will identify the objective function, decision variables, and constraints to write the optimization problem in the usual form. We start by introducing notation based on the problem statement.

Let F be the set of facilities, and let c_f^t be the condition of facility f at the **end**[6] of year t, where t ranges from 1 to 10. Let d_f be the annual deterioration on facility f, and i_f the amount by which the condition will improve if maintenance is performed. So, if no maintenance is performed during year t, then

$$c_f^t = c_f^{t-1} - d_f \qquad \forall f \in F, t \in \{1, 2, \ldots, 10\} \tag{2}$$

and if maintenance is performed we have

$$c_f^t = c_f^{t-1} - d_f + i_f \qquad \forall f \in F, t \in \{1, 2, \ldots, 10\} \tag{3}$$

Both of these cases can be captured in one equation with the following trick: let x_f^t equal one if maintenance is performed on facility f during year t, and 0 if not. Then

$$c_f^t = c_f^{t-1} - d_f + x_f^t i_f \qquad \forall f \in F, t \in \{1, 2, \ldots, 10\} \tag{4}$$

Finally, the condition can never exceed 100 or fall below 0, so the full equation for the evolution of the state is

$$c_f^t = \begin{cases} 100 & \text{if } c_f^{t-1} - d_f + x_f^t i_f > 100 \\ 0 & \text{if} c_f^{t-1} - d_f + x_f^t i_f < 0 \\ c_f^{t-1} - d_f + x_f^t i_f & \text{otherwise} \end{cases} \tag{5}$$

forall $f \in F$ and $t \in \{1, 2, \ldots, 10\}$. Of course, for this to be usable we need to know the current conditions of the facilities, c_f^0.

[6] This word is intentionally emphasized. In this kind of problem it is very easy to get confused about what occurs at the start of period t, at the end of period t, during the middle of period t, etc.

The annual budget can be represented this way: let k_f be the cost of performing maintenance on facility f, and B^t the budget available in year t. Then

$$\sum_{f \in F} k_f x_f^t \leq B^t \qquad \forall t \in \{1, \dots 10\}\,. \tag{6}$$

For the objective, we need the average condition across all facilities and all years; this is simply $\frac{1}{10|F|} \sum_{f\in F} \sum_{t=1}^{10} c_f^t$. The obvious decision variable are the maintenance variables x_f^t, but we also have to include c_f^t because these are influenced by the maintenance variables. As constraints, we need to include the state evolution equations (5), the budget constraints (6), and, less obviously the requirement that x_f^t be either 0 or 1. Putting it all together, we have the optimization problem

$$\begin{aligned}
\min_{\mathbf{x},\mathbf{c}} \quad & \frac{1}{10|F|} \sum_{f \in F} \sum_{t=1}^{10} c_f^t \\
\text{s.t.} \quad & \sum_{f \in F} k_f x_f^t \leq B^t & & \forall t \in \{1, \dots 10\} \\
& c_f^t = \begin{cases} 100 & \text{if } c_f^{t-1} - d_f + x_f^t i_f > 100 \\ 0 & \text{if } c_f^{t-1} - d_f + x_f^t i_f < 0 \\ c_f^{t-1} - d_f + x_f^t i_f & \text{otherwise} \end{cases} & & \forall f \in F, t \in \{1, 2, \dots, 10\} \\
& x_f^t \in \{0, 1\} & & \forall f \in F, t \in \{1, 2, \dots, 10\}
\end{aligned}$$

■

In this example, pay close attention to the use of formulas like (6), which show up very frequently in optimization. It is important to make sure that every "index" variable in the formula is accounted for in some way. Equation (6) involves the variables x_f^t, but for which facilities f and time periods t? The facility index f is summed over, while the time index t is shown at right as $\forall t \in \{1, \dots 10\}$. This means that a copy of (6) exists for each time period, and in each of these copies, the left-hand side involves a sum over all facilities at that time. Therefore the one line (6) actually includes 10 constraints, one for each year. It is common to forget to include things like $\forall t \in \{1, \dots 10\}$, or to try to use an index of summation outside of the sum as well (e.g., an expression like $B_f - \sum_{f \in F} k_f x_f^t$), which is meaningless. Make sure that all of your indices are properly accounted for!

In this next example, the objective function is less obvious.

Example 4. (A facility location problem.) *In a city with a grid network, you need to decide where to locate three bus terminals. Building the terminal at different locations costs a different amount of money. Knowing the home locations of customers throughout the city who want to use the bus service, where should the terminals be located to minimize the construction cost and walking distance customers have to walk? Assume that each customer will walk from their home location to the nearest terminal.*

Solution. This problem could easily become very complicated if we take into account the impact of terminal locations on bus routes, so let's focus on what the problem is asking for: simply locating terminals to minimize walking distance from customers' home locations.

Number each of the intersections in the grid network (Figure 1) from 1 to I, the total number of intersections. Assume that the terminals will be located at these intersections, and let the variables

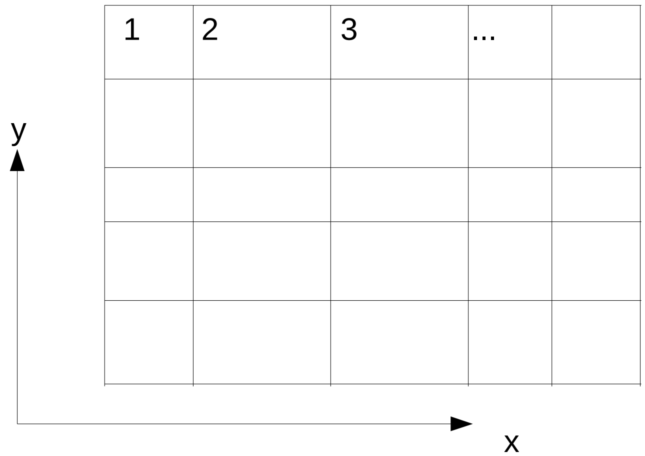

Figure 1: Coordinates and labeling of intersections for Example 4

L_1, L_2, and L_3 denote the numbers of the intersections where terminals will be built. Let $C(i)$ be the cost of building a terminal at location i, so the total cost of construction is $C(L_1)+C(L_2)+C(L_3)$. Let P be the set of customers, and let H_p denote the intersection that is the home location of customer p.

How can we calculate the walking distance between two intersections (say, i and j)? Figure 1 shows a coordinate system superimposed on the grid. Let $x(i)$ and $y(i)$ be the coordinates of intersection i in this system. Then the walking distance between points i and j is

$$d(i,j) = |x(i) - x(j)| + |y(i) - y(j)| \,. \tag{7}$$

This is often called the *Manhattan distance* between two points, after one of the densest grid networks in the world.

So what is the walking distance $D(p)$ for customer p? The distance from p to the first terminal is $d(H_p, L_1)$, to the second terminal is $d(H_p, L_2)$, and to the third is $d(H_p, L_3)$. The passenger will walk to whichever is closest, so $D(p, L_1, L_2, L_3) = \min\{d(H_p, L_1), d(H_p, L_2), d(H_p, L_3)\}$ and the total walking distance is $\sum_{p \in P} D(p, L_1, L_2, L_3)$.

For this problem, the decision variables and constraints are straightforward: the only decision variables are L_1, L_2, and L_3 and the only constraint is that these need to be integers between 1 and I. The tricky part is the objective function: we are instructed both to minimize total cost as well as total walking distance. We have equations for each of these, but we can only have one objective function. In these cases, it is common to form a *convex combination* of the two objectives, introducing a weighting parameter $\Theta \in [0, 1]$. That is, let

$$f(L_1, L_2, L_3) = \Theta[C(L_1) + C(L_2) + C(L_3)] + (1 - \Theta) \left[\sum_{p \in P} D(p, L_1, L_2, L_3) \right] \tag{8}$$

Look at what happens as Θ varies. If $\Theta = 1$, then the objective function reduces to simply minimizing the cost of construction. If $\Theta = 0$, the objective function is simply minimizing the total

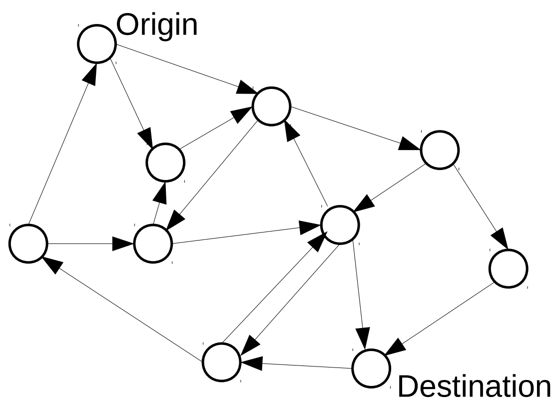

Figure 2: Roadway network for Example 5

walking distance. For a value in between 0 and 1, the objective function is a weighted combination of these two objectives, where Θ indicates how important the cost of construction is relative to the walking distance.

For concreteness, the optimization problem is

$$\begin{aligned} \min_{L_1,L_2,L_3} \quad & \Theta[C(L_1) + C(L_2) + C(L_3)] + (1-\Theta)\left[\sum_{p \in P} D(p, L_1, L_2, L_3)\right] \\ \text{s.t.} \quad & L_f \in \{1, 2, \ldots, I\} \qquad \forall f \in \{1,2,3\} \end{aligned}$$

■

In the final example in this section, finding a mathematical representation of a solution is more challenging.

Example 5. (Shortest path problem.) *Figure 2 shows a road network, with the origin and destination marked. Given the travel time on each roadway link, what is the fastest route connecting the origin to the destination?*

Solution. Notation first: number the intersections 1 to I, and let r and s represent the origin and destination intersections. Number the roadway links from 1 to A, and let t_a be the travel time on link a. So far, so good, but how do we represent a route connecting two intersections?

Following Example 3, introduce binary variables $x_a \in \{0, 1\}$, where $x_a = 1$ if link a is part of the route, and $x_a = 0$ if link a is not part of the route. The travel time of a route is simply the sums of the travel times of the links in the route, which is $\sum_{a=1}^{A} t_a x_a$.

we now have an objective function and decision variables, but what of the constraints? Besides the trivial constraint $x_a \in \{0, 1\}$, we need constraints which require that the x_a values actually form a contiguous path which starts at the origin r and ends at the destination s. We do this by introducing a *flow conservation constraint* at each intersection. For an intersection i, let $F(i)$

denote the set of roadway links which leave intersection i, and let $R(i)$ denote the set of roadway links which enter intersection i.

Consider any contiguous path connecting intersection r and s, and examine any node i. There are one of four cases:

Case I: Node i does not lie on the path at all. Then all of the x_a values should be zero for $a \in F(i) \cup R(i)$.

Case II: Node i lies on the path, but is not the origin or destination. Then $x_a = 1$ for exactly one $a \in F(i)$, and for exactly one $a \in R(i)$.

Case III: Node i is the origin r. Then all x_a values should be zero for $a \in R(i)$, and $x_a = 1$ for exactly one $a \in F(i)$.

Case IV: Node i is the destination s. Then all x_a values should be zero for $a \in F(i)$, and $x_a = 1$ for exactly one $a \in R(i)$.

An elegant way to combine these cases is to look at the differences $\sum_{a \in F(i)} x_a - \sum_{a \in R(i)} x_a$. For cases I and II, this difference will be 0; for case III, the difference will be $+1$, and for case IV, the difference will be -1.

So, this leads us to the optimization formulation of the shortest path problem:

$$
\begin{aligned}
\min_{\mathbf{x}} \quad & \sum_{a \in A} t_a x_a \\
\text{s.t.} \quad & \sum_{a \in F(i)} x_a - \sum_{a \in R(i)} x_a = \begin{cases} 1 & \text{if i = r} \\ -1 & \text{if i = s} \\ 0 & \text{otherwise} \end{cases} & \forall i \in \{1, \ldots, I\} \\
& x_a \in \{0, 1\} & \forall a \in \{1, \ldots, A\}
\end{aligned}
$$

A careful reader will notice that if the four cases are satisfied for a solution, then the equations (3) are satisfied, but the reverse may not be true. Can you see why, and is that a problem? ■

There is often more than one way to formulate a problem: for instance, we might choose to minimize congestion by spending money on capacity improvements, subject to a budget constraint. Or, we might try to minimize the amount of money spent, subject to a maximum acceptable limit on congestion. Choosing the "correct" formulation in this case may be based on which of the two constraints is harder to adjust (is the budget constraint more fixed, or the upper limit on congestion?). Still, you may be troubled by this seeming imprecision. This is one way in which modeling is "as much art as science," which is not surprising — since the human and political decision-making processes optimization tries to formalize, along with the inherent value judgements (what truly is the objective?) are not as precise as they seem on the surface. One hallmark of a mature practitioner of optimization is a flexibility to different formulations of the same underlying problem, and a willingness to engage in "back-and-forth" with the decision maker as together you identify the best formulation for a particular scenario.[7]

[7] As a technical note, in some cases it may not matter. The theory of duality (which is beyond the scope of this course) shows that these alternate formulations can lead to the same ultimate decision, which is comforting.

4 Some Definitions and Useful Transformations

This section shows how we can discuss optimization problems in general, when we don't want to refer to a specific problem context but instead make blanket statements about larger classes of problems. The most generic specification for an optimization problem is to denote all the decision variables with the vector $\mathbf{x}$ (whose dimension is equal to the number of decision variables), the objective function as $f(\mathbf{x})$, and the set of all decision variables which satisfy all the constraints as X. This latter set is called the *feasible set* or *feasible region*, and a vector $\mathbf{x} \in X$ is called a *feasible solution* or simply *feasible.* A feasible solution $\mathbf{x}^* \in X$ is said to be a *global minimum* of $f(\mathbf{x})$ if $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all feasible $\mathbf{x}$, and a *global maximum* of $f(\mathbf{x})$ if $f(\mathbf{x}^*) \geq f(\mathbf{x})$. A global minimum for a minimization problem, or a global maximum for a maximization problem, are called *optimal solutions* or simply *optima.* In general, optima do not always exist; and if they exist they may not be unique. Can you think of some of these cases?

The following results are very useful in showing cases when two superficially different optimization problems may in fact be the same. The first result shows that it is easy to convert any maximization problem to a minimization problem (or vice versa) by negating the objective function. The second shows that constants may be freely added or subtracted to objective functions without changing the optimal solutions. The third shows that multiplication or division by a positive constant does not change the optimal solutions.

Proposition 1. *If the feasible set is X, the solution $\mathbf{x}^*$ be a global maximum of f if and only if $\mathbf{x}^*$ is a global minimum of $-f$ for the same feasible set.*

Proof. For the first part, assume that $\mathbf{x}^*$ is the global maximum of f on the set X. Then $f(\mathbf{x}^*) \geq f(\mathbf{x})$ for all $x \in X$. Multiplying both sides by -1 reverses the sign of the inequality, giving $-f(\mathbf{x}^*) \leq -f(\mathbf{x})$ for all X, so $\mathbf{x}^*$ is a global minimum for $-f$. For the second part, assume that $\mathbf{x}^*$ is the global minimum of $-f$, so $-f(\mathbf{x}^*) \leq -f(\mathbf{x})$ for all $x \in X$. Again multiplying both sides by -1 gives $f(\mathbf{x}^*) \geq f(\mathbf{x})$ for all $x \in X$, so $\mathbf{x}^*$ is a global maximum of f. □

This result is convenient, because we do not need to develop two different results for maximization and minimization problems. Instead, we can focus just on one. Customarily we usually develop results for minimization problems, a convention adopted in this course and fairly widely in engineering. Whenever you encounter a maximization problem, you can convert it to a minimization problem by negating the objective function, then using the minimization results and procedures. (Of course, this choice is completely arbitrary; we could just as well have chosen to develop results only for maximization problems and in fact some other fields do just this.)

Proposition 2. *Let $\mathbf{x}^*$ be an optimal solution when the objective function is $f(\mathbf{x})$ and the feasible set is X. Then for any constant b not depending on $\mathbf{x}$, $\mathbf{x}^*$ is also an optimal solution when the objective function is $f(\mathbf{x}) + b$ and the feasible set is X.*

Proof. By Proposition 1, with no loss of generality we can assume that the optimization problem is a minimization problem and that $\mathbf{x}^*$ is a global minimum. Therefore $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $x \in X$. We can add the constant b to both sides of the equation, so $f(\mathbf{x}^*) + b \leq f(\mathbf{x}) + b$ for all $xinX$. Therefore $\mathbf{x}^*$ is also a global minimum (and thus optimal) when the objective function is $f(\mathbf{x}) + b$. □

Proposition 3. *Let* $\mathbf{x}^*$ *be an optimal solution when the objective function is* $f(\mathbf{x})$ *and the feasible set is* X*. Then for any constant* $c > 0$ *not depending on* $\mathbf{x}$*,* $\mathbf{x}^*$ *is also an optimal solution when the objective function is* $cf(\mathbf{x})$ *and the feasible set is* X*.*

Proof. See exercises. □