Week 12-13: Optimization and root finding.*

Sergei S. Pilyugin†

1 Introduction

The notion of optimization is one of the most important concepts in mathematical applications. Optimization is essentially finding minima or maxima of a given function $f(x)$ (called the *objective function*) where the argument x is restricted to some *feasible set* D.

Example 1. suppose you have a roll of fence of total length L. Using this fence, you may wish to enclose a rectangular region of maximal area A. If x and y denote the dimensions of the rectangle, then the area can be expressed as $A = xy$. Naturally, the variables x and y are not independent since $2x + 2y = L$. Therefore, we can express $y = L/2 - x$ and $A = x(L/2 - x)$. The problem of finding the maximal area is now translated into the problem of finding the maximum value of the function $f(x) = x(L/2 - x)$ on an appropriate feasible set D. The only constraints are that both dimensions of the rectangle be nonnegative, i.e. $x \geq 0$ and $L/2 - x \geq 0$. Hence the feasible set is the closed interval $D = [0, L/2]$. The mathematical notation for the corresponding maximization problem is to find

$$\max_{x \in D} f(x).$$

Definition. (a) A point x^* is a global maximum of f on the set D if $f(x) \leq f(x^*)$ for all $x \in D$. Similarly, a point x^* is a global minimum of f on the set D if $f(x) \geq f(x^*)$ for all $x \in D$.
(b) A point x^* is a local maximum of f in D if $f(x) \leq f(x^*)$ for all $x \in B \cap D$ where B is a sufficiently small ball containing the point x^*. Similarly, a point x^* is a local minimum of f in D if $f(x) \geq f(x^*)$ for all $x \in B \cap D$ where B is a sufficiently small ball containing the point x^*.

A fundamental mathematical result that lays the foundation for solving many optimization problems is the Theorem of Weierstrass. In its simplest form, this Theorem states that any continuous function $f(x)$ defined on a closed and bounded interval D attains its global maximum/minimum on the interval D.

For example, the function $f(x) = x$ attains neither maximum nor minimum on the open interval $(0, 1)$. Yet, if the interval is closed, that is if we include the endpoints $x = 0$ and $x = 1$, the function $f(x) = x$ will attain its global minimum at $x = 0$ and global maximum at $x = 1$.

So far, we have discussed the optimization problems involving continuous functions of continuous arguments defined on some fairly simple feasible sets. This class of problems is typically referred to as unconstrained continuous optimization. There are other important types of optimization problems. For instance, if the feasible set consists of a finite (or discrete) number of points, the corresponding optimization problem is called finite/discrete.

†This course is made possible by the financial support from the Howard Hughes Medical Institute.

Example 2. A classical example of a finite optimization problem is the so-called backpack problem. Suppose we have a finite number of items $X = \{x_1, x_2, ..., x_n\}$ where each x_i is a nonnegative value that denotes the space occupied by the corresponding item. In addition, we have a backpack of some given capacity $C > 0$. Now, we wish to find a way to pack the backpack to the maximal possible capacity. In mathematical terms, we want to find the maximal possible sum of distinct elements of the set X provided that such sum does not exceed the total capacity C. On the surface, this problem is rather trivial, since we may simply try all possible combinations of items from X, and find the optimal packing combination. But the problem is that the total number of combinations of different combinations of items is given by 2^n (the number of all subsets of the set with n elements), and it grows very fast (exponentially fast) with n. Hence the *heuristic approach* involving a complete combinatorial search quickly becomes computationally prohibitive.

2 Critical Points and First Variation Principle

The general method for finding minima and maxima for smooth functions $f(x)$ involves finding *critical points* and comparing the values at the critical points to the values on the boundary of the feasible region.

What is a critical point of a function? First lets consider a differentiable function $f(x)$ defined on some interval D. Let x^* be a point of local maximum/minimum in the interior of the interval D. For definiteness, lets consider a local maximum. On the one hand, for all sufficiently values $t > 0$, we have $f(x + t) \leq f(x^*)$, hence

$$f'(x^*) = \lim_{t \to 0^+} \frac{f(x+t) - f(x^*)}{t} \leq 0.$$

On the other hand, for all sufficiently values $t > 0$, we have $f(x - t) \leq f(x^*)$, hence

$$f'(x^*) = \lim_{t \to 0^+} \frac{f(x-t) - f(x^*)}{-t} \geq 0.$$

Therefore, we have $0 \leq f'(x^*) \leq 0 \Rightarrow f'(x^*) = 0$. Same holds if x^* is an interior local minimum. We conclude that the only interior points that may be candidates for maxima/minima are the points x^* such that $f'(x^*) = 0$. These are called the critical points of f inside the interval D.

To generalize the concept of a critical point to functions of several variables, let us imagine an abstract observer that moves inside the feasible region D. Suppose that the function $f(\mathbf{x})$ attains a maximum at some interior point $\mathbf{x}^* \in D$. A point is interior to the region D if it is contained in some sufficiently small ball that lies entirely inside the region D. Let the observer move along some smooth parametric curve $\mathbf{x}(t)$ such that $\mathbf{x}(0) = \mathbf{x}^*$. From the point of view of the observer, the observed values of the function f are given parametrically by $f(\mathbf{x}(t))$. Regardless of the chosen path, the observed function $f(\mathbf{x}(t))$ will have a local maximum at $t = 0$. Therefore, it must hold that

$$\frac{d}{dt} f(\mathbf{x}(0)) = 0.$$

Using the Chain Rule for multivariate functions, this condition may be rewritten as

$$\frac{\partial f(\mathbf{x}^*)}{\partial x_1} x_1'(0) + \frac{\partial f(\mathbf{x}^*)}{\partial x_2} x_2'(0) + ... + \frac{\partial f(\mathbf{x}^*)}{\partial x_n} x_n'(0) = 0.$$

Since the path of the observer is arbitrary, so are the values $x_i'(0),\ i = 1, 2, .., n$. Consequently, it must be the case that

$$\frac{\partial f(\mathbf{x}^*)}{\partial x_1} = \frac{\partial f(\mathbf{x}^*)}{\partial x_2} = ... = \frac{\partial f(\mathbf{x}^*)}{\partial x_n} = 0.$$

Using the vector notation, we form the vector containing all partial derivatives of f

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n} \right),$$

called the *gradient* of the function f. Now we can simply define a critical point $\mathbf{x}^*$ as a point where the gradient is zero, $\nabla f(\mathbf{x}^*) = \mathbf{0}$.

Example 1 (cont-d). Empowered by these concepts, let us finish the problem with the maximal rectangular area enclosed by a fence of length L. Recall, that we need to find the maximum of the objective $f(x) = x(L/2 - x)$ on the feasible interval $[0, L/2]$. Solving

$$f'(x) = L/2 - 2x = 0 \Rightarrow x^* = L/4,$$

we find a unique interior critical point. Comparing the value $f(L/4) = L^2/16$ to the values at the endpoints $f(0) = f(L/2) = 0$, we find that the point $x^* = L/4$ is the point of global maximum on $[0, L/2]$. Hence, the rectangle of the maximal area must be a square such that $x^* = y^* = L/4$.

Now, let us apply this method to a multivariate function.

Example 3. Suppose that we wish to construct a rectangular box that maximizes volume for a given surface area A. Let x, y, z be the dimensions of the box. Then the surface area can be expressed as

$$2(xy + xz + yz) = A \Rightarrow z = \frac{A/2 - xy}{x + y} = z(x, y).$$

The volume of the box now can be expressed as

$$V = f(x, y) = x\, y\, z(x, y) = \frac{xy}{x + y}(A/2 - xy).$$

The feasible region D is defined as follows

$$D = \{(x, y) | x, y \geq 0,\ xy \leq A/2\}.$$

The objective function f is zero everywhere on the boundary of the feasible region D, so to find a maximum, we will locate the critical point(s) of f in the interior of D. Computing the partial derivatives, we find that

$$\frac{\partial f}{\partial x} = \frac{y^2(A/2 - xy) - xy^2(x + y)}{(x + y)^2}, \quad \frac{\partial f}{\partial y} = \frac{x^2(A/2 - xy) - yx^2(x + y)}{(x + y)^2}.$$

Setting both of these expressions equal to zero, and simplifying the resulting system, we obtain the conditions for the critical point(s):

$$A/2 - xy = x(x + y) = y(x + y).$$

Solving this system, we find a unique critical point $x^* = y^* = \sqrt{A/6}$. Hence, the maximal value of f is given by $f(x^*, y^*) = (A/6)^{\frac{3}{2}}$. The resulting box with the maximal volume is a cube with all three dimensions given by

$$x^* = y^* = z^* = \sqrt{A/6}.$$

Exercises. Use methods of this section to solve the following minimax problems:
1. Find the dimensions of the rectangular box (without the lid) that maximizes the volume with a fixed surface area. Remark: the surface area is a sum of five rectangular areas $A = xy + 2xz + 2yz$.
2. Find the dimensions of the rectangular box that minimizes the surface area with a fixed volume. Are you really surprised by the answer?

3 Constrained optimization. Lagrange multipliers

So far, we have discussed the unconstrained optimization problems. What does it make if we impose an additional *constraint* on the feasible region D and what is considered a constraint? Suppose we still wish to find

$$\max_{\mathbf{x}\in D} f(\mathbf{x}),$$

but the set of all feasible arguments D must also satisfy an additional condition $g(\mathbf{x}) = 0$. If g is a nondegenerate function, that is $\nabla g \neq \mathbf{0}$, we call the condition $g(\mathbf{x}) = 0$ a *regular constraint*.

Let us follow up with the moving observer analogy. Suppose that $\mathbf{x}^*$ is a point of local maxima/minima of the constrained problem. Let $\mathbf{x}(t)$ be a path of the observer such that $\mathbf{x}(0) = \mathbf{x}^*$ and $g(\mathbf{x}(t)) = 0$ for all values of t. Essentially, the observer is no longer allowed to move freely, because the motion is now confined to the surface defined by the constraint $g = 0$. Differentiating $f(\mathbf{x}(t))$ at $t = 0$, we find that

$$\frac{d}{dt} f(\mathbf{x}(0)) = \nabla f(\mathbf{x}^*) \cdot \mathbf{x}'(0) = 0,$$

because $\mathbf{x}^*$ is a point of local maxima/minima. Differentiating $g(\mathbf{x}(t))$ at $t = 0$, we find that

$$\frac{d}{dt} g(\mathbf{x}(0)) = \nabla g(\mathbf{x}^*) \cdot \mathbf{x}'(0) = 0,$$

because $g(\mathbf{x}(t)) \equiv 0$. In other words, for any vector $\mathbf{x}'(0)$ which satisfies $\nabla g(\mathbf{x}^*) \cdot \mathbf{x}'(0) = 0$, we must also have that $\nabla f(\mathbf{x}^*) \cdot \mathbf{x}'(0) = 0$. Using the orthogonality concept, we may say that $\nabla f(\mathbf{x}^*)$ must be orthogonal to any vector that is in turn orthogonal to $\nabla g(\mathbf{x}^*)$. The only case in which this can happen is when the two vectors $\nabla f(\mathbf{x}^*)$ and $\nabla g(\mathbf{x}^*)$ are collinear (i.e., parallel). Since $\nabla g(\mathbf{x}^*) \neq 0$, there must exist a scalar λ such that

$$\nabla f(\mathbf{x}^*) = \lambda \nabla g(\mathbf{x}^*).$$

Such scalar is called the *Lagrange multiplier* for the constrained problem. We conclude that each critical point of the constrained problem must admit the corresponding Lagrange multiplier. Multipliers may (and generally will) differ for different critical points.

Let us apply the concept of the Lagrange multiplier to the fence problem. This time, we will treat the area $A = f(x, y) = xy$ as the objective function of two variables subject to the constraint $g(x, y) = 2x + 2y - L = 0$. The gradients of the objective function and the constraint are given by

$$\nabla f = (y, x), \quad \nabla g = (2, 2),$$

respectively. If λ is the Lagrange multiplier, then we must have

$$(y^*, x^*) = \lambda(2, 2) \Rightarrow x^* = y^* = 2\lambda.$$

Solving the constraint equation for x^* and y^*, we find that $x^* = y^* = L/4$.

A similar solution exists for the rectangular box of maximal volume. Here, we have the objective function $f(x, y, z) = xyz$ subject to the constraint $g(x, y, z) = 2(xy + xz + yz) - A = 0$. Computing the gradients and introducing the Lagrange multiplier, we obtain the system

$$\nabla f = (yz, xz, xy) = 2\lambda(y + z, x + z, x + y) = \lambda \nabla g.$$

A point of maximum volume clearly cannot correspond to $\lambda = 0$, thus $\lambda \neq 0$ and $x, y, z \neq 0$. After some algebraic manipulation, we obtain

$$\frac{1}{2\lambda} = \frac{1}{y} + \frac{1}{z} = \frac{1}{x} + \frac{1}{z} = \frac{1}{x} + \frac{1}{y},$$

hence at a critical point (x^*, y^*, z^*), we have

$$\frac{1}{x^*} = \frac{1}{y^*} = \frac{1}{z^*} \Rightarrow x^* = y^* = z^* = \sqrt{A/6}.$$

4 Linear Programming

5 Root finding

6 MATLAB notes

References