Constrained Optimization

ML Tutorial

Shreeya Venneti

Agenda

- Convex function - formal definition
- What is the convex optimization problem?
- Equality constraints
- Inequality constraints
- KKT conditions

Convex Set

- A convex set is a set in a vector space that satisfies the following property:

  For any two points, x and y, that belong to the set, and for any value of t between 0 and 1 ($0 \leq t \leq 1$), the point $tx + (1 - t)y$ also belongs to the set.

  - In more formal terms, a set S is convex if, for all x, y in S and for all t in [0, 1], the combination **$tx + (1 - t)y$** is also in S.

  - This definition essentially means that a convex set contains all the line segments connecting any two points within the set.

  - This definition will be used to define convex functions further.

Convex Function

- A function f(x) is considered convex if, for any two points x and y in its domain, and for any value of t between 0 and 1 ($0 \le t \le 1$), the following inequality holds:

$$\mathbf{f(tx + (1 - t)y) \le tf(x) + (1 - t)f(y)}$$

- In other words, a function is convex if the line segment connecting any two points on its graph always lies above the graph.

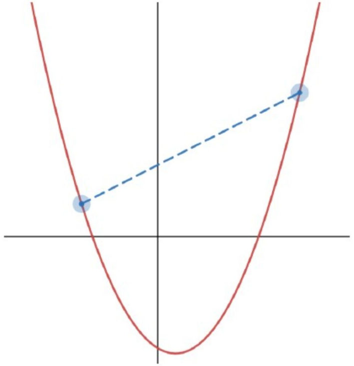

Figure 1: Visualization of convexity

Why are we interested in convex functions?

- Convex functions have a special property of having a single minimum - i.e local minimum is the global minimum

- Why so? Imagine a function with multiple peaks and trenches, i.e multiple local minima

- Draw a line connecting two points on either side of a peak

- Violates the convex function principle - there would be points on the curve above the line

Convex Optimization Problem

- A convex optimization problem is a mathematical optimization problem where the goal is to minimize (or sometimes maximize) a convex objective function subject to a set of convex constraints.

- The general form of a convex optimization problem can be expressed as follows:

- Minimize (or Maximize) f(x) subject to:
  - **$g_i(x) \leq 0$ for $i = 1, 2, \ldots, m$**
  - **$h_j(x) = 0$ for $j = 1, 2, \ldots, p$**

- In this formulation:
  - f(x) is the objective function, and it is convex. If the problem is a maximization problem, then -f(x) is usually taken instead
  - **$g_i(x)$** are convex inequality constraints, and they are convex functions. These constraints define regions in the feasible space.
  - **$h_j(x)$** are convex equality constraints, and they are typically affine (linear) functions.

Some jargon

- $f(x)$ - objective function
- g_i - inequality constraints ($i = \{1, \ldots, m\}$)
- h_j - equality constraints ($j = \{1, \ldots, p\}$)
- Feasible set - set of points where all the constraints are satisfied

Equality Constraint

- Let f(x,y) be the function to be minimized

- Let h(x,y) = 0 be the equality constraint

- We construct the Lagrangian as follows:

$$\mathbf{L(x,y,\lambda) = f(x,y) + \lambda h(x,y)}$$

- The point (x,y) where f is minimum and h(x,y) = 0 is the point where h(x,y) is a tangent to f(x,y) at the minima

- Lagrangian is a recipe to solve convex optimization problems and is nothing but the cost associated with violating the constraints

Equality Constraint

- Consider two equality constraints: $h_1(x,y,z) = 0$ & $h_2(x,y,z) = 0$
- The solution point will lie on both the planes simultaneously and on the curve f
- Lagrangian here would be: $\mathbf{L(x, y, z, \lambda_1, \lambda_2) = f(x,y) + \lambda_1 h_1(x,y) + \lambda_2 h_2(x,y)}$
- To generalise: $\mathbf{L(\bar{x},\lambda) = f(\bar{x}) + \Sigma\lambda_i h_i(\bar{x})}$
- Practice: $f(x) = x^2 + y^2 + z^2$; $h_1 = x + y = 3$; $h_2 = x - y = 3$

The KKT conditions

1. Stationarity Condition
   a. This condition states that the gradient of the Lagrange function with respect to the each of the decision variables should be zero.

$$\boldsymbol{\nabla_{\bar{x},\lambda}(x,\lambda) = 0}$$

2. Primal Feasibility
   a. Inequality constraints to be satisfied

$$\mathbf{g_i(x) \leq 0 \text{ for all } i = 1, 2, \ldots, m}$$

3. Dual Feasibility
   a. The lagrange multipliers should be non-negative

$$\boldsymbol{\lambda_j >= 0 \text{ for all } j = 1,2,\ldots p}$$

The KKT conditions

4. Complementary Slackness

   a. Complementary Slackness captures the relationship between the Lagrange multipliers and the constraints

   b. The complementary slackness condition tells us that if a constraint is not binding i.e., there is slack ($\mathbf{g_j(x) < 0)}$, then the corresponding Lagrange multiplier must be zero. If the Lagrange multiplier is not zero, it implies that the constraint is active i.e ($\mathbf{g_j(x) = 0)}$ and the solution lies on the constraint boundary.

   c. Either the constraint is active ($g_i(x) = 0$), in which case the Lagrange multiplier (λ_i) should be positive, or if the Lagrange multiplier is zero, the constraint is not active

$$\boldsymbol{\lambda_j\, g_j(x) = 0 \text{ for all } j = 1,2,\ldots,m}$$

Inequality Constraints

- Just like with equality constraints, in the case of inequality constraints, we introduce Lagrange multipliers.

- Combine the objective function and the inequality constraints with lagrange multipliers

- Apply KKT conditions

- Practice:

    - Minimize the objective function: $f(x,y) = x^2+y^2$ subject to the inequality constraints: $x+y-1 \leq 0$ and $x-y+2 \leq 0$
    - Minimize the objective function: $f(x) = x^2$ subject to the inequality constraint: $g(x) = x - 2 \leq 0$

Practice

- Take cases:
  - Case 1: Both constraints active
  - Case 2: 1st active, 2nd inactive
  - Case 3: 1st inactive, 2nd active
  - Case 4: both inactive

- For each case construct the lagrangian and check the validity of the KKT conditions and consistency with other constraints