Tutorial 4: Optimization

Mohammadreza Safavi
mohammadreza.safavi@mail.utoronto.ca

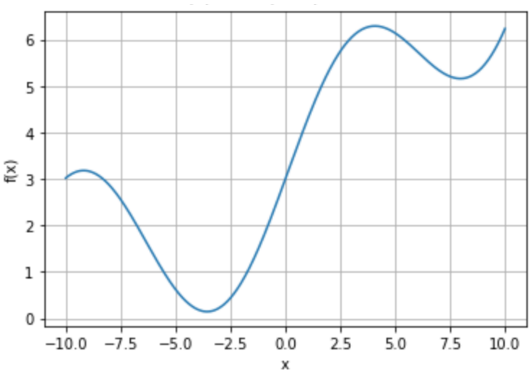
University of Toronto
ECE1513: Introduction to Machine Learning
Instructor: Dr. Ali Bereyhi

February 3, 2025

Problem of Optimization

The optimization problem can be formulated as:

$$p^* = \min_x f(x) \quad \text{(e.g., loss function)}$$



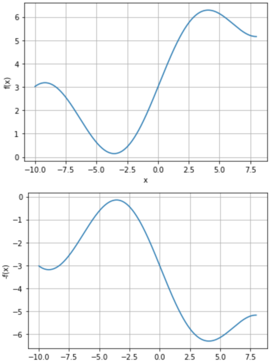
Problem of Optimization

Sometimes, we deal with maximization problems:

$$p^* = \max_x f(x) \quad \text{(e.g., reward function)}$$

This can be converted to a minimization problem by multiplying by -1:

$$p^* = -\min_x(-f(x))$$

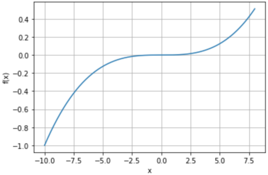


Local and Global Minima

Some functions have multiple minima:

- The lowest value is called the **global minimum**.
- Other minima are called **local minima**, as they are only the minimum in their neighborhood.

A function may also have **saddle points** where the gradient is zero but is neither a minimum nor a maximum.



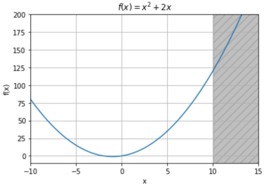
Optimization with Constraints

In real-world problems, we often have constraints:

$$p^* = \min_x f(x) \quad \text{s.t. } h(x) \leq 0$$

Example:

$$\begin{aligned} \min_x \quad & x^2 + 2x \\ \text{s.t.} \quad & x \geq 10 \end{aligned}$$



The standard derivative approach may not work when constraints are present.

Optimization with Constraints

In the general case, we can have multiple constraints:

$$
\begin{aligned}
p^* = \min_{\mathbf{x}} \quad & f_o(\mathbf{x}) \\
\text{s.t.} \quad & f_i(\mathbf{x}) \leq 0, \ i = 1, ..., m \\
& h_i(\mathbf{x}) = 0, \ i = 1, ..., n
\end{aligned}
$$

The standard derivative approach may not work when constraints are present.

Lagrangian Formulation

When constraints are involved, we define the Lagrangian:

$$\mathcal{L}(\mathbf{x}, \lambda, \nu) = f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{n} \nu_i h_i(\mathbf{x})$$

In case of only having inequality constraints:

$$\mathcal{L}(\mathbf{x}, \lambda) = f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x})$$

The optimal solution satisfies:

$$p^* = \min_{\mathbf{x}} \max_{\lambda \geq 0} \mathcal{L}(\mathbf{x}, \lambda)$$

Lagrangian Formulation

In case of only having inequality constraints:

$$\mathcal{L}(\mathbf{x}, \lambda) = f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x})$$

The optimal solution satisfies:

$$p^* = \min_{\mathbf{x}} \max_{\lambda \geq 0} \mathcal{L}(\mathbf{x}, \lambda)$$

Now define the dual problem:

$$d^* = \max_{\lambda \geq 0} \min_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \lambda)$$

Under certain conditions (e.g., Slater's condition), strong duality holds: $d^* = p^*$.

Lagrangian Formulation

So if those conditions hold, we can solve the dual problem instead of the primal problem.
If f_0, f_i, h_i are convex and differentiable and we have Slater's condition (i.e., there exists a feasible point), then the optimal solution can be found by solving the dual problem.
If the primal problem is convex, then the dual problem is also convex and we can take the derivative of the dual function to find the optimal solution.
Let's define another function called the dual function:

$$\phi(\lambda) = \min_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \lambda)$$

Lagrangian Formulation

Under certain conditions we can use the dual function to find the optimal solution:

$$\phi(\lambda) = \min_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \lambda)$$

$$p^* = d^* = \max_{\lambda \geq 0} \phi(\lambda)$$

Lagrangian Formulation

In summary, under certain conditions (e.g., convexity, Slater's condition), we can use the dual function to find the optimal solution.
First we need to define the Lagrangian, then the dual function, and finally solve the dual problem.

$$\phi(\lambda) = \min_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \lambda)$$

$$p^* = d^* = \max_{\lambda \geq 0} \phi(\lambda)$$

Example: Constrained Optimization

Solve the problem:

$$\min_{x,y,z} x^2 + y^2 + 2z^2 \quad \text{s.t. } 2x + 2y - 4z \geq 8$$

Example: Constrained Optimization

Solve the problem:

$$\min_{x,y,z} x^2 + y^2 + 2z^2 \quad \text{s.t. } 2x + 2y - 4z \geq 8$$

Define the Lagrangian:

$$\mathcal{L}(x, y, z, \lambda) = x^2 + y^2 + 2z^2 + \lambda(8 - 2x - 2y + 4z)$$

Define the dual function:

$$\phi(\lambda) = \min_{x,y,z} \mathcal{L}(x, y, z, \lambda)$$

Example: Constrained Optimization

$$\phi(\lambda) = \min_{x,y,z} \mathcal{L}(x, y, z, \lambda)$$

$$\phi(\lambda) = \min_{x,y,z} x^2 + y^2 + 2z^2 + \lambda(8 - 2x - 2y + 4z)$$

$$\frac{\partial \mathcal{L}}{\partial x} = 2x - 2\lambda = 0 \rightarrow x = \lambda$$

$$\frac{\partial \mathcal{L}}{\partial y} = 2y - 2\lambda = 0 \rightarrow y = \lambda$$

$$\frac{\partial \mathcal{L}}{\partial z} = 4z + 4\lambda = 0 \rightarrow z = -\lambda$$

Example: Constrained Optimization

$$\phi(\lambda) = \min_{x,y,z} \mathcal{L}(x, y, z, \lambda) = \mathcal{L}(x = \lambda, y = \lambda, z = -\lambda, \lambda)$$

$$\phi(\lambda) = \lambda^2 + \lambda^2 + 2\lambda^2 + \lambda(8 - 2\lambda - 2\lambda + 4(-\lambda))$$

$$\phi(\lambda) = 8\lambda - 4\lambda^2$$

Example: Constrained Optimization

$$\phi(\lambda) = 8\lambda - 4\lambda^2$$

Find the optimal value of λ to solve the dual problem:

$$\frac{d\phi}{d\lambda} = 8 - 8\lambda = 0 \rightarrow \lambda^* = 1$$

Final Solution

$$\lambda^* = 1, \quad x^* = 1, \quad y^* = 1, \quad z^* = -1$$

The optimal function value is $p^* = 4$.

Visualizing the Solution

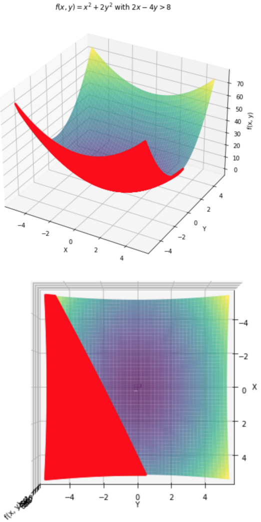
For a better understanding, let's visualize the solution but for a simpler case:

$$\min_{x,y} x^2 + 2y^2 \quad \text{s.t. } 2x - 4y \geq 8$$

After solving the problem, we can find that $x^* = 4/3, y^* = -4/3$ and the optimal value is $p^* = 16/3$.

Visualizing the Solution

Lets visualize the problem:



Visualizing the Solution

Now see the solution: