Convex Optimization Problem

What concepts should we know first?

Optimization

Unconstrained example: gradient descent

\( \min f(x) \)

\( x^* \) is the optimal solution to

\( f(x) \) and \( \nabla f(x^*) = 0 \)

Constraint

\( \min f(x) \)

\( \text{s.t. } h_i(x) = 0 (i = 1, \ldots, m) \)

Equality constraint

we minimize our objective w.r.t. some inequality and equality constraint

\( f(x) \) objective function

Can be a cost, energy function, or reward

\( \max x = (x_1, \ldots, x_n) \) n variables

in the objective function

In a space if we consider two points \( x_1, x_2 \) and try to connect them two ways:

\( x = \theta x_1 + (1-\theta) x_2 \)

we can consider it as an affine set \( a^T x = b \) general format

contains the line through any two points in the set

linear equation

Convex Set

Convex Set, non-convex set

Convex Set is a set that considering any two points in the set, the line segment between the two points is in the set!
Convex Function

\[ f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha)f(x_2) \]

where \( 0 \leq \alpha \leq 1 \)

Jensen's Inequality / Zero-order Convexity

\[ f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha)f(x_2) \]

Another way to look at a convex function

First-Order Convexity

If \( f(x) \) is differentiable

\[ f(x) \geq f(x_0) + \nabla f(x_0)^T (x-x_0) \]

Example: \( f(x) = x^2 \)

\[ x^2 \geq x_0^2 + 2x_0(x-x_0) \Rightarrow x^2 \geq x_0^2 + 2x_0x - 2x_0^2 \Rightarrow x^2 \geq 2x_0x - x_0^2 \]

\[ x^2 + x_0^2 - 2x_0x \geq 0 \Rightarrow (x-x_0)^2 \geq 0 \]

\( f(x) \) is a convex function
Second-order convexity

If \( f(x) \) is twice-differentiable

\[ \nabla^2 f(x) \rightarrow f(x) \text{ is a convex function} \]

Hessian

Now, we can go back to our original question! ?!

What is a convex optimization problem? !

Convex Optimization

**Unconstrained**

\[ \min f_0(x) \]

\( x^* \) is an optimal solution.

\[ \nabla f_0(x^*) = 0 \]

If \( f_0(x) \) is a convex function

**Constraint**

\[ \min f_0(x) \]

\[ \text{s.t. } f_i(x) \leq 0 \quad i = 1, \ldots, m \text{ inequality} \]

\[ h_i(x) = 0 \quad i = 1, \ldots, P \text{ equality} \]

\( \text{If } f_0(x) \text{ is a convex function} \)

\( \text{and } f_i(x) \text{ is convex sets} \)

Many problems can be solved via convex optimization.

It's reliable and efficient - can be solved in polynomial time.

Any local min of \( f_0(x) \) that is feasible is also a global optimal!

How to solve a convex problem? (unconstrained convex optimization)

Lagrangian \( \Pi \) is a function that combines both constraints and objective terms.

\[ \min f_0(x) \]

\[ \text{s.t. } f_i(x) \leq 0 \quad i = 1, \ldots, m \]

\[ h_i(x) = 0 \quad i = 1, \ldots, P \]

\[ \Pi : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \]

\[ \Pi (x, \lambda, \mu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{P} \mu_i h_i(x) \]

\( \lambda, \mu \) are Lagrange multipliers

\( \lambda \) - Lagrange multiplier for inequality

\( \mu \) - Lagrange multiplier for equality

\( \lambda > 0 \quad \text{for inequality} \)

\( \mu \geq 0 \quad \text{for equality constraints} \)
\[ f_0(x) = 1 - x_1^2 - x_2^2 \]
\[ h(x) = x_1 + x_2 - 1 = 0 \]
\[ \lambda(x, y) = 1 - x_1^2 - x_2^2 + \lambda (x_1 + x_2 - 1) \]

\[ \frac{\partial \lambda}{\partial x_1} = -2x_1 + \lambda = 0 \implies x_1 = \frac{-\lambda}{2} \implies x_1 = \frac{1}{2} \]

\[ \frac{\partial \lambda}{\partial x_2} = -2x_2 + \lambda = 0 \implies x_2 = \frac{\lambda}{2} \]

Optimal Solution
\[ \left( \frac{1}{2}, \frac{1}{2} \right) \]

we have \[ x_1 + x_2 - 1 = 0 \implies \frac{1}{2} + \frac{1}{2} - 1 = 0 \implies \lambda = 1 \]

\[ f_0(x): \text{primal problem} \]
\[ \min x_1^2 + 1 \]
\[ \text{St. } (x - 2)(x - 4) \leq 0 \implies x^2 - 6x + 8 \leq 0 \]

\[ \lambda(x, \lambda) = f_0(x) + \lambda h(x) = x_1^2 + 1 + \lambda(x_1 + x_2 - 1) = (\lambda + 1)x_1^2 + 6\lambda x + 8\lambda + 1 \]

\[ \frac{\partial \lambda}{\partial x} = 2(\lambda + 1)x - 6\lambda = 0 \implies x^* = \frac{3\lambda}{\lambda + 1} \]

Plugging the \( x^* \) into the Lagrangian

\[ (\lambda + 1) \left( \frac{3\lambda}{\lambda + 1} \right)^2 - 6\lambda \left( \frac{3\lambda}{\lambda + 1} \right) + 8\lambda + 1 = \frac{9\lambda^2}{\lambda + 1} - \frac{18\lambda^2}{\lambda + 1} + 8\lambda + 1 \]
\[ - \frac{9\lambda^2}{\lambda + 1} + 8\lambda + 1 \rightarrow \begin{cases} - \frac{9\lambda^2}{\lambda + 1} + 8\lambda + 1 & \lambda > 0 \\ - \alpha & \lambda \leq -1 \end{cases} \]
Now, we have a new function in terms of $\lambda$, which is not a convex function. (Concave)

But if we can maximize this new function, provides the lower bound to the solution of the main problem (which we call it primal problem).

This new function is called Lagrangian dual.

$$g(\lambda) = \max \frac{-9 \lambda^2}{\lambda + 1} + 8 \lambda + 1$$  \hspace{1cm} \text{dual problem}

st \hspace{0.5cm} \lambda \geq 0

much easier to solve

$$\frac{\partial g}{\partial \lambda} = \frac{-18 \lambda (\lambda + 1) - 9 \lambda^2}{(\lambda + 1)^2} + 8 = 0 \Rightarrow \lambda^* = 2$$

$$\hat{x} = \frac{3 \times 2}{(2+1)} = 2 \quad \hat{x}^* = 2$$

$$x^2 + 1 = (2)^2 + 1 = 5$$

The optimal solution is $$(2, 5)$$

lower bound property $d^* \lambda \geq 0$ $g(\lambda) \leq p^*$

d$^*$ is the solution to the dual $\implies d^* = p^*$ $\implies$ Strong duality holds.
Here are the actual plots for the second example. Showing the objective function in blue, constrains in red. The black lines in the next figure is the Lagrangian function with different lambda values.

![Feasible Area](image)

Figure 1: Feasibility area and objective function
For lambda value equal to zero, you have your original objective function and as you increase the lambda you may change your plot a little bit but that does not affect the minimum point.

Figure 2: Feasibility area, objective function and Lagrangian functions
The Dual function is maximized in (2,5) and gives the optimal solution to the dual problem as well as primal (original) problem.

Figure 3: $g(\lambda)$