Problemas de otimização

Ricardo Brauner

31 Julho, 2020

Instituto UFC Virtual

HUAWEI

Optimization Problem

- Find x that minimizes or maximizes function f(x).

$$min(.max)\, f(x)$$

$$s.t. \quad g_i(x) \geq 0, i = 1,2,\ldots,m, \text{ inequality constraints}$$

$$h_j(x) = 0, j = 1,2,\ldots,p, \text{ equality constraints}$$

where $x = (x_1, x_2, \ldots, x_n)$.

- Objective or utility function when maximizing
- Cost, loss or error function when minimizing

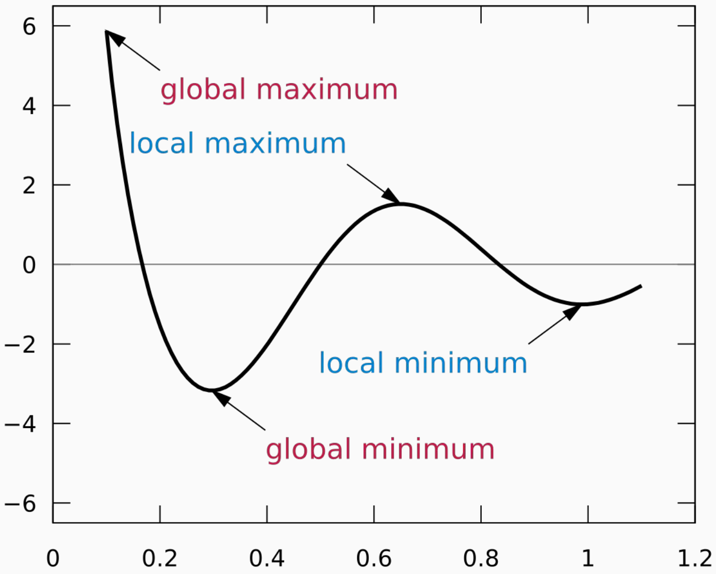
Classification of Optimization Problems

With no constraints, it can be expressed as:

$$\min f(x)$$

The common method is Fermat theorem.

If $f'(x) = 0$, the critical point is obtained. Then, verify that the extreme value can be obtained at the critical point.



Constraint optimization

- Sometimes, the maximized or minimized $f(x)$ function under all possible values is not what we desire.
- Instead, we might want to find the maximum or minimum value of $f(x)$ when x is in a certain collection s.
- The points within the collection s are called feasible points.

Equality constraints

With equality constraints, it can be expressed as:

$$\min f(x)$$
$$s.t.\quad h_i(x) = 0, i = 1,2,\ldots,n.$$

The common method is Lagrange multiplier method, that is, introducing n Lagrange multipliers λ to construct Lagrange function

$$L(x,\lambda) = f(x) + \sum_{i=1}^{n} \lambda_i h_i(x)$$

and then seeking the partial derivative of each variable to be zero. With the collection of candidate values we can get the optimal value through verification.

Inequality constraints

With inequality constraints, it can be expressed as:

$$\min f(x)$$
$$s.t.\quad h_i(x) = 0, i = 1,2,\ldots,n,$$
$$g_j(x) \le 0, j = 1,2,\ldots,m.$$

A common method is to introduce new variables λ_i and μ_j, to Generalized Lagrangian functions based on all equality, inequality constraints and f (x).

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{n} \lambda_i h_i(x) + \sum_{j=1}^{m} \mu_j g_j(x)$$

KKT (kuhn-kuhn-tucker) conditions

- A common method is to introduce new variables λ_i and μ_j, to Generalized Lagrangian functions based on all equality, inequality constraints and $f(x)$.

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{n} \lambda_i h_i(x) + \sum_{j=1}^{m} \mu_j g_j(x)$$

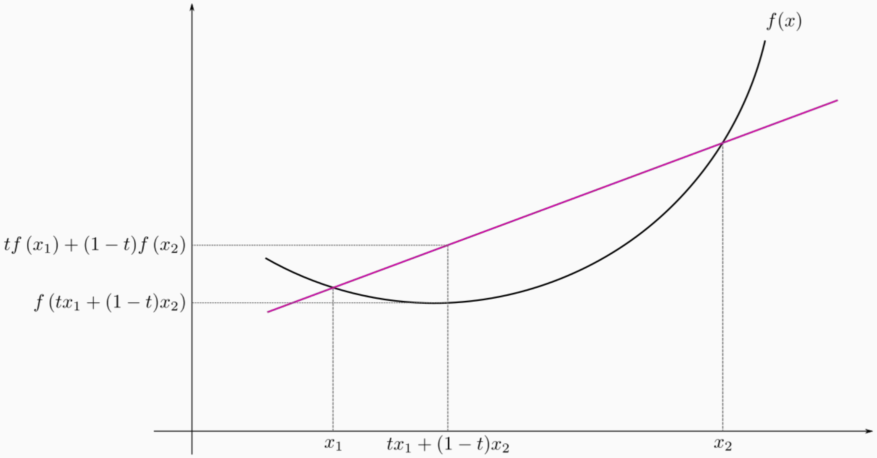
- Not sufficient
- Depends on the type of problem

Convex function

For $\lambda \in (0,1)$, given arbitrary a, b

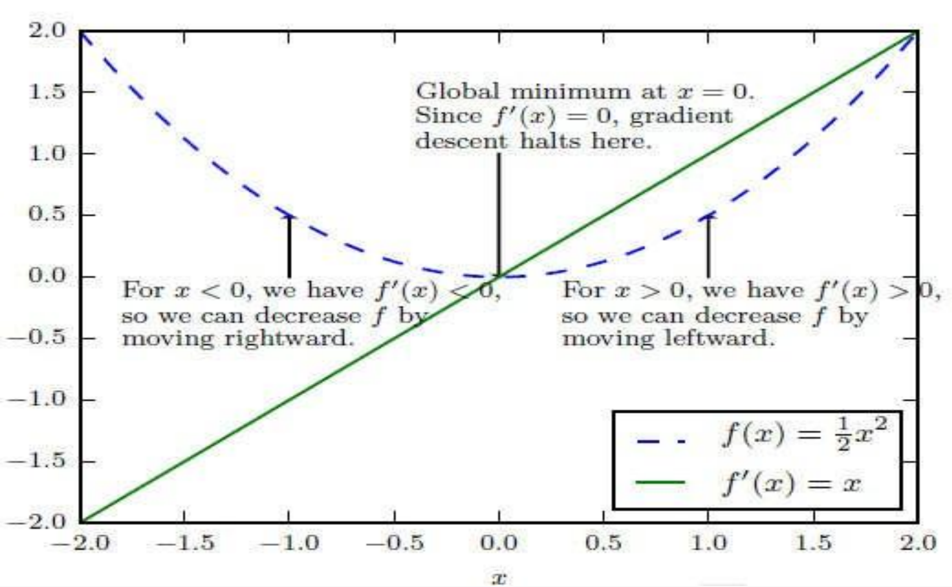
$f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda)f(b)$

The extremum point of the convex function is present at the stationary point.



Gradient Based Optimization Method

- The derivative indicates how to change x to slightly improve y.
- Move x in the opposite direction of the derivative by a small step to reduce f(x).



Gradient

To the case of multidimensional functions, the partial derivative is used to describe the degree of variation of the function relative to the respective variable.

Gradient is the vector whose components are the partial derivatives of f(x) at p represented as ∇f(x).

$$\nabla f(p) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(p) \\ \vdots \\ \frac{\partial f}{\partial x_n}(p) \end{bmatrix}$$

Gradient

The derivative of f x) in the direction of u (unit vector) is $u^T \nabla f(x)$.

For a task to minimize $f(x)$, we want to find the direction with the fastest downward change.

$$\min u^T \nabla f(x) = \min \|u\|_2 \|\nabla f(x)\|_2 \cos\theta$$

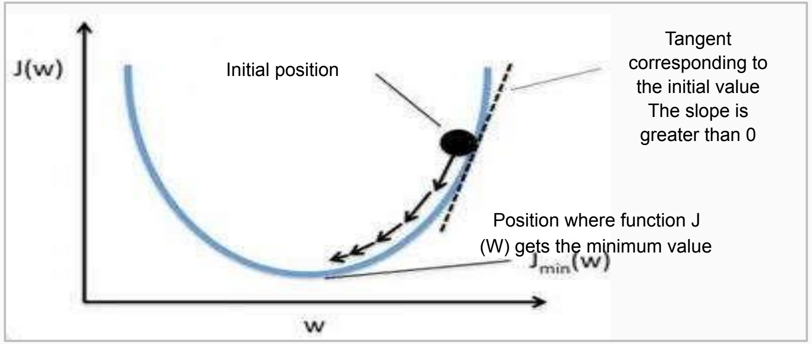
You can see that the direction in which f (x) value decreases the maximum is the negative direction of the gradient.

Steepest descent

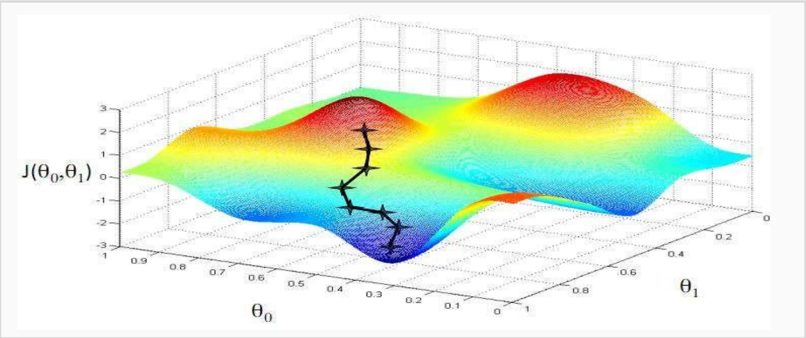
- A negative gradient vector points downhill.
- Steepest descent or gradient descent.
- Under the gradient descent method, the update point is proposed as:

$$x' = x - \varepsilon \nabla f(x)$$

- ε is the learning rate.
- Converges when the gradient is zero or approaching zero.



Two-dimensional space



Multi-dimensional space