Lagrange Multipliers

Young Shen

December 12, 2024

1 Equality Constraint

Considering:

$$\max_x \quad f(x,y) \tag{1.1}$$
$$\text{s.t.} \quad h_i(x,y)=0, \quad i=1,\dots,n \tag{1.2}$$

This is an optimization problem with n equality constraints. As we learnt in Calculus class, using **Lagrange Multiplier** will turn this into an optimization problem with $n+1$ parameters with no constraint.

$$\mathcal{L}(x,\lambda)=f(x,y)+\sum_{i=1}^{n}\lambda_i h_i(x,y), \quad \lambda_i \neq 0 \tag{1.3}$$

then take the derivative to x and λ's, we get a **feasible solution**:

$$\begin{aligned} &\nabla\mathcal{L}(x,y)=0 \\ &\frac{\partial \mathcal{L}}{\partial \lambda_i}=0, \quad i=1,\dots,n \end{aligned} \tag{1.4}$$

To understand why this works, consider the problem in 2 dimension with only one constraint first (the arrows show the direction where f(x,y) grows):

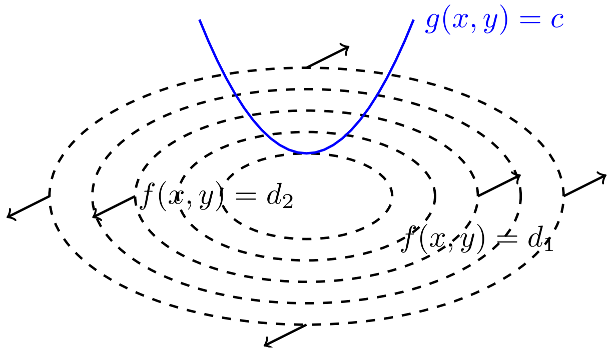

Figure 1: Contours of f and g

from this figure1, we can see the feasible solution is where blue line is **tangent** to the dashed line. At this point, the gradient of f and g is **parallel**:

$$\nabla f=\lambda\nabla g, \quad \lambda\neq 0 \tag{1.5}$$

which is exactly the first line in (1.4). Note that the second line is equivalent to the constraint.

2 Non-equality constraint

Considering:

$$\max_x \quad f(x,y) \tag{2.1}$$
$$\text{s.t.} \quad h_i(x,y)=0, \quad g_j(x,y)\leq 0, \quad i=1,\dots,n, \quad j=1,\dots,m \tag{2.2}$$

To make the understanding easier, consider the 2D situation where there is only one g constraint. For g, there are 2 possible conditions (**suppose the field $g \leq 0$ is the upper side of blue curve**):

- the optimal point is in its range 2
- the optimal point is out of its range 3

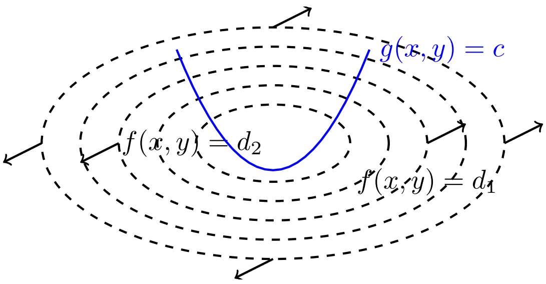

Figure 2: Optimal point **in** its range

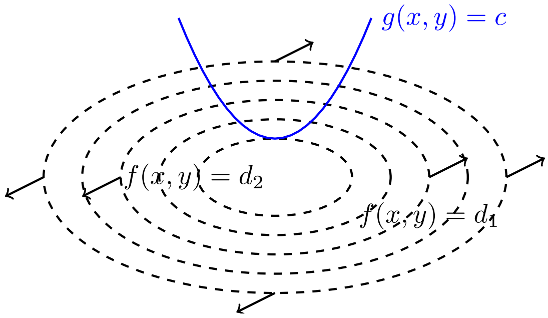

Figure 3: optimal point **out** of its range

In the first case, the constraint of g has no contribute, and the optimal is reached when $g \leq 0$:

$$\begin{aligned} &\mu = 0, \quad \text{(the constraint of } g \text{ has no contribute)} \\ &g(x^*, y^*) < 0, \quad \text{(optimal point is in its range)} \end{aligned} \tag{2.3}$$

where (x^*, y^*) is the optimal point. And then we directly optimize f:

$$\nabla \mathcal{L}(x, y) = 0 \tag{2.4}$$

In the second case, the optimal point is reached when $g = 0$, here the constraint of g becomes equality constraint, we treat g as h. Don't forget the assumption: **the field $g \leq 0$ is the upper side of blue curve**, which indicate that, **at the optimal point, the gradient of g and f are pointing opposite direction** :

$$\nabla f = -\mu \nabla g, \quad \mu > 0 \tag{2.5}$$

Similarly, construct the Lagrangian:

$$\mathcal{L}(x, \lambda, \mu) = f(x, y) + \mu g(x, y), \quad \mu > 0 \tag{2.6}$$

and then let:

$$\begin{aligned} &\nabla \mathcal{L}(x, y) = 0 \\ &\frac{\partial \mathcal{L}}{\partial \mu} = 0 \\ &\mu > 0 \end{aligned} \tag{2.7}$$

and then solve for x^*.

Note that we can combine the 2 cases mentioned before:

$$\begin{aligned} \nabla \mathcal{L}(x, y) &= 0 \\ g(x^*, y^*) &\leq 0 \\ \mu &\geq 0 \\ \mu g(x^*, y^*) &= 0 \end{aligned} \tag{2.8}$$

This result can be generalized to higher dimensions with more constraints.

$$\begin{aligned} \nabla \mathcal{L} &= 0 \\ \frac{\partial \mathcal{L}}{\partial \lambda_i} &= 0, \quad i = 1, \ldots, n \\ \mu_j &\geq 0, \quad j = 1, \ldots, m \\ g_j &\leq 0, \quad j = 1, \ldots, m \\ \mu_j g_j &= 0, \quad j = 1, \ldots, m \end{aligned} \tag{2.9}$$

This is called **Karush-Kuhn-Tucker(KKT) conditions**.