Optimization under constraint

Math 251

October 15, 2015

1 Optimization under one constraint

1.1 In two dimensions

Let $f(x, y)$ be given. Previously we discussed how to find the max and min of $f(x, y)$ over the plane and over a region $\mathbb{R}$. There is yet another type of optimization: finding the extrema of f along a curve. That is we consider

$$\max_{(x,y)} f(x, y)$$
$$\text{s.t. } g(x, y) = c.$$

We say $f(x, y)$ is the objective function and $g(x, y) = c$ is the constraint that the maximization of f is subject to. We seek a necessary condition that describes the candidates for the optimal points (a, b). Note that $g(x, y) = c$ describes a curve so it effectively reduces our optimization problem to 1 dimension. That is let $\mathbf{r}(t)$ be a parametrization corresponding to the curve $g(x, y)$ then the problem becomes

$$\max_{t} f(\mathbf{r}(t)).$$

But here we can just differentiate in t and use the first derivative condition:

$$f'(t) = \nabla f(\mathbf{r}(t)) \cdot \mathbf{r}'(t) = 0. \tag{1}$$

On the other hand, $\mathbf{r}(t)$ is a parametrization corresponding to the curve $g(x, y)$ means that

$$g(\mathbf{r}(t)) = c.$$

Thus

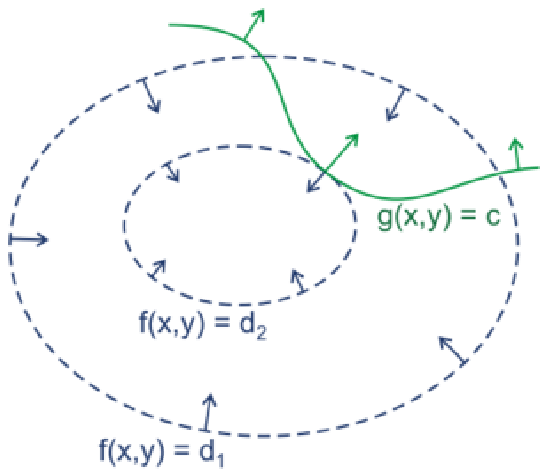
$$\frac{d}{dt}g(\mathbf{r}(t)) = \nabla g(\mathbf{r}(t)) \cdot \mathbf{r}'(t) = 0. \tag{2}$$

Comparing (1) and (2) we see that (since we are in 2 dimensions) ∇f must be parallel to ∇g at the point t_0 on the curve that satisfies (1) and (2). That is

$$\nabla f(a, b) = \lambda \nabla g(a, b) \tag{3}$$

if (a, b) is the optimal point. λ is called the Lagrange multiplier of the optimization problem.

The equation (3) has a nice geometrical interpretation:



to optimize f along the curve $g(x, y) = c$, we consider various level curves of f. At the point of intersection of these two curves, if the gradient of f is not orthogonal to the tangent of the curve $g(x, y)$ we can move along the curve $g(x, y) = c$ to arrive at a higher level curve of f. Thus we can repeat this process until the gradient of f is orthogonal to the tangent of $g(x, y)$. This is the point where the maximum may happen (because we will move out of the curve $g(x, y) = c$ if we follow the gradient of f at this point).

1.2 In three dimensions

Now consider the function $f(x, y, z)$. We want to optimize f subject to the constraint:

$$g(x, y, z) = c.$$

The technique is exactly the same as above. We let $\mathbf{r}(s,t)$ be a parametrization of the surface given by $g(x,y,z) = c$. This reduces f as a function of 2 variables s,t. Thus the first order condition requires

$$\frac{\partial}{\partial s} f(\mathbf{r}(s,t)) = \frac{\partial}{\partial t} f(\mathbf{r}(s,t)) = 0.$$

But this becomes

$$\nabla f \cdot \frac{\partial}{\partial s}\mathbf{r}(s,t) = 0 \quad (4)$$

$$\nabla f \cdot \frac{\partial}{\partial t}\mathbf{r}(s,t) = 0. \quad (5)$$

We can verify that since $g(\mathbf{r}(s,t)) = c$, g satisfies exactly the equation (4). Thus ∇f and ∇g must be parallel again. Thus there must exists a λ so that

$$\nabla f = \lambda \nabla g.$$

2 Optimization under multiple constraints

Now consider the function $f(x,y,z)$. We want to optimize f subject to two constraints:

$$\begin{aligned} g_1(x,y,z) &= c_1 \\ g_2(x,y,z) &= c_2. \end{aligned}$$

Let $\mathbf{r}(t)$ be the parametrization of the curve that satisfies

$$\begin{aligned} g_1(\mathbf{r}(t)) &= c_1 \\ g_2(\mathbf{r}(t)) &= c_2. \end{aligned}$$

Then we require

$$\frac{d}{dt} f(\mathbf{r}(t)) = \nabla f \cdot \mathbf{r}'(t) = 0. \quad (6)$$

On the other hand

$$\nabla g_1 \cdot r'(t) = 0 \quad (7)$$

$$\nabla g_2 \cdot r'(t) = 0. \quad (8)$$

Assuming $\nabla g_1, \nabla g_2$ are not parallel, then (6) and (7) say that ∇f must be in the plane determined by $\nabla g_1, \nabla g_2$. That is there must exist λ, μ so that

$$\nabla f = \lambda \nabla g_1 + \mu \nabla g_2.$$