§18

18.1 Constrained Optimization

The following is the idea behind constrained optimization:

> A local extremum of $f\colon \mathcal{D}^n \to \mathbf{R}$ constrained to $S \subseteq \mathcal{D}^n$ must occur at a point $\boldsymbol{a}$ in S such that $\nabla f(\boldsymbol{a})$ is orthogonal to S at $\boldsymbol{a}$.

Orthogonal to S at $\boldsymbol{a}$ means orthogonal to the tangent space of S at $\boldsymbol{a}$. We need f to have continuous partial derivatives and S to be "nice" near $\boldsymbol{a}$. The intuition is that $\nabla f(\boldsymbol{a})$ orthogonal to S at $\boldsymbol{a}$ means that along S the function f is constant near $\boldsymbol{a}$, which is the condition for an extremum.

The practical form is called "Lagrange Multipliers":

Theorem 1. Consider functions $g_1, \ldots, g_k\colon \mathcal{D}^n \to \mathbf{R}$ with continuous partial derivatives, and define $S \subseteq \mathcal{D}^n$ be an intersection of level sets $g_1^{-1}(c_1) \cap \cdots \cap g_k^{-1}(c_k)$. If $\boldsymbol{a}$ in S is a local extremum of $f\colon \mathcal{D}^n \to \mathbf{R}$ constrained to S and $\nabla g_1(\boldsymbol{a}), \ldots, \nabla g_k(\boldsymbol{a})$ are linearly independent, then

$$\nabla f(\boldsymbol{a}) = \lambda_1 \nabla g_1(\boldsymbol{a}) + \cdots + \lambda_k \nabla g_k(\boldsymbol{a})$$

for some λ_i in $\mathbf{R}$.

Note 1. The $\lambda_1, \ldots, \lambda_k$ are called Lagrange multipliers. The practical form follows from the boxed statement because $\nabla g_1(\boldsymbol{a}), \ldots, \nabla g_1(\boldsymbol{a})$ span the set of vectors perpendicular to S at $\boldsymbol{a}$.

Here is how to apply the theorem. To optimize $f\colon \mathcal{D}^n \to \mathbf{R}$ subject to the constraints

$$\begin{cases} g_1(x_1, \ldots, x_n) = c_1 \\ \quad\vdots \qquad\qquad \vdots \\ g_k(x_1, \ldots, x_n) = c_k \end{cases}$$

define a new function:

$$\begin{aligned} &F(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_k) \\ &= f(x_1, \ldots, x_n) - \lambda_1[g_1(x_1, \ldots, x_n) - c_1] - \cdots - \lambda_1[g_k(x_1, \ldots, x_n) - c_k] \end{aligned}$$

Then find the unconstrained local extrema of F. If $(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_k)$ is an unconstrained local extrema of F, then $(x_1, \ldots, x_n)$ is a constrained local extrema of f. All constrained local extrema of f may be obtained in this way, except possibly those $(x_1, \ldots, x_n)$ for which

$$\nabla g_1(x_1, \ldots, x_n), \ldots, \nabla g_k(x_1, \ldots, x_n)$$

are dependent. To find the unconstrained local extrema of F, we require $\nabla F = \mathbf{0}$, which says:

$$\begin{cases} \nabla f(x_1, \ldots, x_n) = \lambda_1 \nabla g_1(x_1, \ldots, x_n) + \cdots + \lambda_k \nabla g_k(x_1, \ldots, x_n) \\ \qquad\qquad c_1 = g_1(x_1, \ldots, x_n) \\ \qquad\qquad \vdots \qquad\quad \vdots \\ \qquad\qquad c_k = g_k(x_1, \ldots, x_n) \end{cases}$$

One may start with these equations instead of creating the auxiliary function F, but it is good to know both methods.

Example 1. Find the absolute extrema of $f(x, y) = 2x + 3y$ constrained to:

$$S = \{(x, y) \in \mathbf{R}^2 \mid x^2 + 4y^2 = 16\}$$

Solution. Define

$$F(x, y, \lambda) = f(x, y) - \lambda(x^2 + 4y^2 - 16) = 2x + 3y - \lambda(x^2 + 4y^2 - 16)$$

so $\nabla F = \mathbf{0}$ means:

$$\begin{cases} 2 = \lambda(2x) \\ 3 = \lambda(8y) \\ 16 = x^2 + 4y^2 \end{cases}$$

Rewrite the third equation as

$$16\lambda^2 = (x\lambda)^2 + 4(y\lambda)^2$$

but $x\lambda = 1$ and $y\lambda = 3/8$ so $16\lambda^2 = 25/16$. Therefore $\lambda^2 = 25/16^2$ and hence $\lambda = \pm 5/16$. Then the first two equations give the corresponding values for x and y, yielding the solutions:

$$(x, y, \lambda) = (-16/5, -6/5, -5/16)$$
$$(x, y, \lambda) = (16/5, 6/5, 5/16)$$

The respective values of f at these values are -10 and 10. Therefore f constrained to S has minimum value -10 at $(-16/5, -6/5)$ and maximum value 10 at $(16/5, 6/5)$. ∎

Example 2. Find the absolute extrema of $f(x,y,z) = x$ constrained to:

$$S = \{(x,y,z) \in \mathbf{R}^3 \mid y^2 + z^2 = 1 \text{ and } x + y - z = 0\}$$

Solution. Define

$$\begin{aligned} F(x,y,z,\lambda_1,\lambda_2) &= f(x,y,z) - \lambda_1(y^2 + z^2 - 1) - \lambda_2(x + y - z) \\ &= x - \lambda_1(y^2 + z^2 - 1) - \lambda_2(x + y - z) \end{aligned}$$

so $\nabla F = \mathbf{0}$ means:

$$\begin{cases} 1 = \lambda_1(0) + \lambda_2(1) \\ 0 = \lambda_1(2y) + \lambda_2(1) \\ 0 = \lambda_1(2z) + \lambda_2(-1) \\ 1 = y^2 + z^2 \\ 0 = x + y - z \end{cases}$$

The first equation gives $\lambda_2 = 1$. Rewrite the fourth equation as

$$\lambda_1^2 = (y\lambda_1)^2 + (z\lambda_1)^2$$

and use $y\lambda_1 = -1/2$ and $z\lambda_1 = 1/2$ to obtain $\lambda_1^2 = 1/2$. Therefore $\lambda_1 = \pm 1/\sqrt{2}$. If $\lambda_1 = 1/\sqrt{2}$, then

$$(x,y,z,\lambda_1,\lambda_2) = (\sqrt{2}, -1/\sqrt{2}, 1/\sqrt{2}, 1/\sqrt{2}, 1)$$

and if $\lambda_1 = -1/\sqrt{2}$, then:

$$(x,y,z,\lambda_1,\lambda_2) = (-\sqrt{2}, 1/\sqrt{2}, -1/\sqrt{2}, -1/\sqrt{2}, 1)$$

The absolute minimum of f is $-\sqrt{2}$ occurring at $(-\sqrt{2}, 1/\sqrt{2}, -1/\sqrt{2})$ and the absolute maximum of f is $\sqrt{2}$ occurring at $(\sqrt{2}, -1/\sqrt{2}, 1/\sqrt{2})$. ■

Example 3 (From Licata 14.2). Find extrema of $f(x,y) = -x^2 - y^2$ on:

$$S = \{(x,y) \in \mathbf{R}^2 \mid 2x - 3y = 1\}$$

Solution. Answer: No minimum. Maximum of f is $-1/13$ occurring at $(2/13, -3/13)$. ■

Example 4 (From Licata 14.3). Find extrema of $f(x,y,z) = 2xy + 2yz + 2xz$ on:

$$S = \{(x,y,z) \in \mathbf{R}^3 \mid xyz = 64\}$$

Solution. Answer: No minimum. No Maximum. ∎

Example 5 (From Licata 14.4). Find extrema of $f(x, y, z) = -2y+5x^2+3z^2$ on:

$$S = \{(x, y, z) \in \mathbf{R}^3 \mid x^2 + y^2 = 4z^2\}$$

Solution. Answer: Minimum $-4/3$ at $(0, 4/3, -2/3)$. No maximum. ∎

Example 6 (From Licata 14.6). Find extrema of $f(x, y, z) = x^2+y^2+z^2$ on:

$$S = \{(x, y, z) \in \mathbf{R}^3 \mid 2x^2 + 3y^2 + 4z^2 = 1 \text{ and } y + z = 0\}$$

Solution. Answer: Minimum $2/7$ at the points $(0, 1/\sqrt{7}, -1/\sqrt{7})$ and $(0, -1/\sqrt{7}, 1/\sqrt{7})$. Maximum $1/2$ at $(1/\sqrt{2}, 0, 0)$ and $(-1/\sqrt{2}, 0, 0)$. ∎

Example 7 (From Licata 14.7). Find extrema of $f(x, y, z) = xy + xz$ on:

$$S = \{(x, y, z) \in \mathbf{R}^3 \mid x - y - 2z = 0 \text{ and } x + y + z = 3\}$$

Solution. Answer: No minimum. Maximum $9/4$ at $(3/2, 3/2, 0)$. ∎