11.8 Lagrange Multipliers

Exploration: Suppose we want to find the extrema of $f(x, y) = x^2 + 2y^2$ but also know that any (x, y) must be on the unit circle. Graphically, $z = f(x, y)$ produces a paraboloid whose vertex is at the origin and opens in the direction of the positive z-axis.

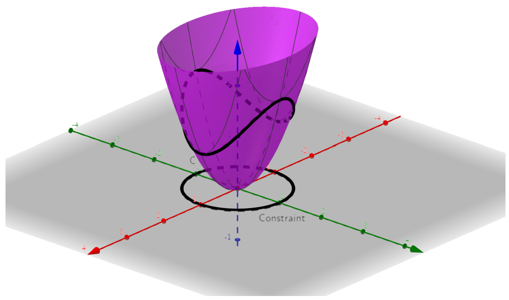

Figure 1: https://www.geogebra.org/3d/kmw4uhnw

Now, to find the maximum of $f(x, y) = x^2 + 2y^2$ subject to $x^2 + y^2 = 1$ is to find the largest value of c such that the level curve $f(x, y) = c$ intersects $x^2 + y^2 = 1$. For convenience, let's call $g(x, y) = x^2+y^2$. Now, this happens specifically when these curves just touch each other; this happens specifically when the curves share a tangent line; this happens specifically when the normal lines are parallel; this happen specifically when $\nabla f(x_0, y_0) = \lambda \nabla g(x, y)$ for some $\lambda \in \mathbb{R}$.

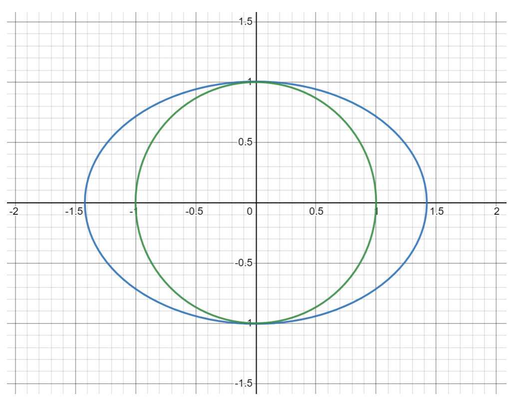

Figure 2: https://www.desmos.com/calculator/wenqh2aldm

To formalize this, suppose a function f has an extremum at $P(x_0, y_0, z_0)$ on the surface S, and let C be a curve with vector equation $\mathbf{r}(t) = \langle x(t), y(t), z(t)\rangle$ that lies on S and passes through P. Further suppose $t_0 \in \mathbb{R}$ such that $\mathbf{r}(t_0) = \langle x_0, y_0, z_0\rangle$. Then $h(t) = f((x(t), y(t), z(t))$ represents the values of f that lie on C. Since f has an extremum at (x_0, y_0, z_0), h must have an extremum at t_0, so by Fermat's Theorem, $h'(t_0) = 0$. Moreover,

$$\begin{aligned} 0 &= h'(t_0) \\ &= f_x(x_0, y_0, z_0)x'(t_0) + f_y(x_0, y_0, z_0)y'(t_0) + f_z(x_0, y_0, z_0)z'(t_0) \\ &= \nabla f(x_0, y_0, z_0) \cdot \mathbf{r}'(t_0) \end{aligned}$$

Thus, $\nabla f(x_0, y_0, z_0)$ and $\mathbf{r}'(t_0)$ must be orthogonal. Since we know that the gradient vector at P is perpendicular to the tangent vector to any curve C on S at P, we know that $\nabla g(x_0, y_0, z_0)$ is also orthogonal to $\mathbf{r}'(t_0)$. Since both $\nabla f(x_0, y_0, z_0)$ and $\nabla g(x_0, y_0, z_0)$ are orthogonal to $\mathbf{r}'(t_0$, it follows that $\nabla f(x_0, y_0, z_0)$ is parallel to $\nabla g(x_0, y_0, z_0)$. Therefore, if $\nabla g(x_0, y_0, z_0) \neq 0$, then there must exist a number λ such that

$$\nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0)$$

Definition

The number λ such that $\nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0)$ is called a **Lagrange multiplier**.

Method of Lagrange Multipliers

To optimize a multivariable function $f(x, y, z)$ subject to the constraint $g(x, y, z) = k$ (assuming extrema exist),

1. Find all values of x, y, z, λ such that

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z) \qquad \text{and} \qquad g(x, y, z) = k$$

2. Evaluate f at all of the point (x, y, z) that you found in the previous step. The largest of these values is the maximum of f, and the least value is the minimum of f.

Note: Solving for x, y, z, λ is not necessarily straightforward. In three variables, solving $\nabla f(x, y, z) = \lambda \nabla g(x, y, z)$ produces four equations in four unknowns. That is,

$$f_x = \lambda g_x \tag{1}$$
$$f_y = \lambda g_y \tag{2}$$
$$f_z = \lambda g_z \tag{3}$$
$$g = k \tag{4}$$

In two variables, solving $\nabla f(x, y) = \lambda \nabla g(x, y)$ produces three equations in three unknowns. That is,

$$f_x = \lambda g_x \tag{1}$$
$$f_y = \lambda g_y \tag{2}$$
$$g = k \tag{3}$$

The kicker is that these are *not necessarily linear systems*. Linear algebra studies solving linear systems, but we very well may have a nonlinear system, so you'll have to use some ingenuity.

Example 1. Find the extrema of $f(x, y) = x^2 + 2y^2$ on the unit circle.

Example 2. Find the extrema of $f(x, y) = x^2 + 2y^2$ on the disk $x^2 + y^2 \leq 1$.

Example 3. Find the points on the sphere $x^2 + y^2 + z^2 = 4$ closest to and furthest from the point $(3, -2, 6)$.