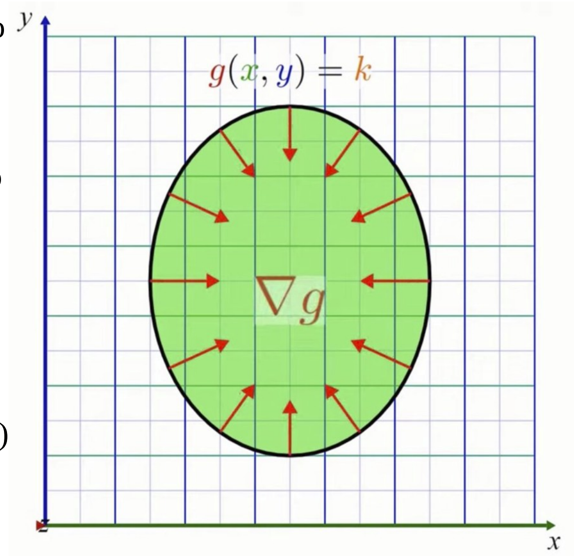
Lagrange multipliers

Suppose we wish to maximize the value of the two-dimensional function $z = f(x, y)$ along within some region. In order to do so we must check for both local maxima within the region as well as points along the border of the region. The latter is analagous to checking for extreme values of $f(x)$ at the end points of an interval except that in the one-dimensional case we only have to check two points; in the multidimensional case, by contrast, there are an infinite number of points on the boundary of the region. We cannot check them all one-by-one, but we can find maximize the function another way. Let us begin by writing the boundary constraint as



$$g(x, y) = k. \tag{1}$$

Since $g(x, y) = k$ lies entirely in the x-y plane, its gradient ∇g also lies in the x-y plane in a direction perpendicular to $g(x, y)$ (see figure).

By contrast, the gradient of $z = f(x, y)$ does not necessarily lie in the x-y plane. However if we only look at the *projection* of $f(x, y)$ onto $g(x, y) = k$, i.e. we consider only the points that lie directly above or below $g(x, y) = k$, then the extreme points on this curve will be those whose gradients when projected onto the x-y plane are *parallel to* ∇g.

In other words, at the extreme points there exists some constant λ (the so-called “Lagrange multiplier”) such that

$$\frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x} \text{ and } \frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial x}.$$

This is equivalent to saying that if we set $f(x, y)$ equal to some *constant* value such that $\dfrac{\partial f}{\partial z} = 0$, then the gradients of f and g are parallel.

$$\nabla f = \lambda \nabla g \tag{2}$$

As a counterpoint, consider what would happen if the gradient of $f(x, y)$ at some point on this projection had a component *perpendicular* to ∇g; in that case, we could follow the curve of $g(x, y) = k$ to some other nearby point on $g(x, y) = k$ and the value of $f(x, y)$ would either increase or decrease. Only if $f(x, y)$ is already *at an extreme* point would moving along the curve not result in a higher or lower point.

To find the extreme values of f along the constraint curve $g(x, y) = k$, we must solve the constraint equation (1) simultaneously with the set of gradient equations (2).

This argument generalizes to functions of n variables except that in the latter case, equation (2) becomes a system of n equations.