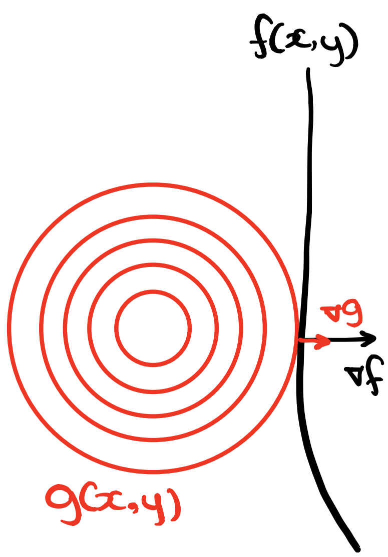
Lagrange multipliers can be used to find the minima or maxima which satisfy some constraint. That constraint is also represented by a function.

The key insight popularized by Lagrange is that the contours of the function must intersect with the contour of the constraint — else the constraint be violated.

In order for the function's contour to be maximal it must be greater than all other contour lines of it's own function. The same can be said of the minimal contour.

The only way for the functions minimum/max to maintain that definition is if the contour is exactly the optimal solution.

Since the function and constraint are continuous without exactness there will always be a slight deviation in variables which are slightly more minimal/maximal.



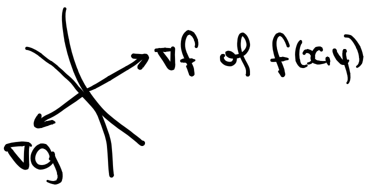
Where $f(x,y)$ is the function we are trying to optimize and $g(x,y)$ is the constraint.

Since the contour lines of $f(x,y)$ for the min/max always intersect the contour $g(x,y)$ their gradients will be parallel.

However those gradients may not be equal because they are distinct functions.

Want to maximize $f(x,y) = x^2y$

Example: constrained to $g(x,y) = x^2+y^2$
$= a^2$, where a is some constant.



Want to solve $\nabla f = \lambda \nabla g$.

where lambda is the lagrange multiplier.

Example: $f(x) = x^2y$

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \qquad \nabla g = \begin{bmatrix} \frac{\partial g}{\partial x} \\ \frac{\partial g}{\partial y} \end{bmatrix}$$

$$= \begin{bmatrix} 2xy \\ x^2 \end{bmatrix} \qquad = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

Solving, $\nabla f = \lambda \nabla g$,

$$\left.\begin{matrix} 2x\lambda \\ 2y\lambda \end{matrix}\right| \begin{matrix} 2xy \\ x^2 \end{matrix}$$

$$\left.\begin{matrix} \lambda \\ 2y\lambda \end{matrix}\right| \begin{matrix} y \\ x^2 \end{matrix}$$, dividing row #1 by $2x$, meaning $\lambda = y$

$$\left.\begin{matrix} \lambda \\ \sqrt{2y\lambda} \end{matrix}\right| \begin{matrix} y \\ x \end{matrix}$$, rooting row #2, meaning

$x \pm \sqrt{2y\lambda}$
$\pm \sqrt{2y^2}$, subbing in $y = \lambda$
$\pm \sqrt{2}y$

Re-express the original function in terms of the f gradients relationship to g-gradient by way of $x^2+y^2 = a^2$ var y. and solve for y in terms

$$= (\sqrt{2}y)^2 + y^2$$

of the divergences from zero (eg a).

$$= 2y^2 + y^2$$

$$a^2 = 3y^2$$

$$y^2 = a^2/3$$

$$y = \sqrt{\frac{a^2}{3}}$$

$$= a/\sqrt{3}$$

The solutions of x can be,

$$\frac{a}{\sqrt{3}}\begin{bmatrix}\sqrt{2}\\1\end{bmatrix}, \frac{a}{\sqrt{3}}\begin{bmatrix}\sqrt{2}\\-1\end{bmatrix}, \frac{a}{\sqrt{3}}\begin{bmatrix}-\sqrt{2}\\1\end{bmatrix}, \frac{a}{\sqrt{3}}\begin{bmatrix}-\sqrt{2}\\-\sqrt{1}\end{bmatrix}$$

Solutions to $f(x,y) = a^2 = x^2 + y^2$,

$$\underset{\text{max}}{\frac{a^3}{3\sqrt{3}}2}, \quad \underset{\text{min}}{\frac{-2a^3}{3\sqrt{3}}}, \quad \underset{\text{max}}{\frac{2a^3}{3\sqrt{3}}}, \quad \underset{\text{min}}{\frac{-2a^3}{3\sqrt{3}}}$$

This is useful when trying to find max/min with fixed relationships between variables in a model – like a constraint!

There are more technicalities and more formal ways of solving the constrained optimization problem.

1. $\mathcal{L}(x,y,\ldots,\lambda) = f(x,y,\ldots) - \lambda(g(x,y,\ldots) - k)$, k is the "finite" constraint
   - ↳ get partial derivatives to $\lambda, x, y, \ldots$
   - ↳ solve simultaneous equations for the vals of each variable and λ (the lagrange mult.)
   - ↳ sub in max/min vars into orig func. $\Rightarrow \nabla\mathcal{L} = \vec{0}$.

(*) The lagrangian multiplier expresses the proportionate relationship between the function and the constraint

$$\lambda_i = \frac{d^{(1)} f(x_i, y_i, \ldots)}{d k_i}$$

. technically there can be multiple solutions/combinations $(x_i, y_i, \ldots)$ satisfying the constrain k_i .