Lagrange Multipliers:

The question Lagrange multipliers help us answer is: What are the maximum and minimum of a function $f(x,y)$ when I'm restricted to the set of points $g(x,y) = 0$?

To give a conceptual explanation of how this leads to the equation $\nabla f(a,b) = \lambda \nabla g(a,b)$, let's first discuss optimization of f without the constraint of $g(x,y) = 0$ in a slightly different way.

So suppose I have some function $f(x,y)$ with gradient ∇f. As an example:

Let $f(x,y) = x^3 + y^3 - 3x - 3y$

$$\nabla f = \langle 3x^2 - 3, 3y^2 - 3 \rangle$$

Now if we set the gradient equal to 0, we get four critical points: $(1,1), (1,-1), (-1,1), (-1,-1)$

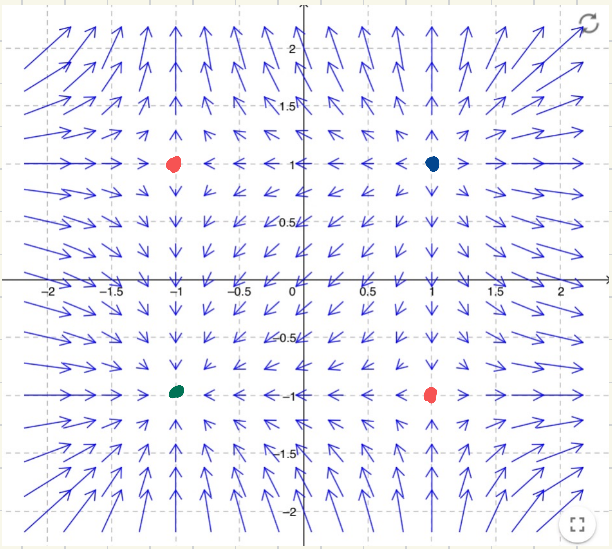
The discriminant is $36xy$, so our four pts are classified as:

$(1,1)$: Local min

$(1,-1)$: Saddle point

$(-1,1)$: Saddle point

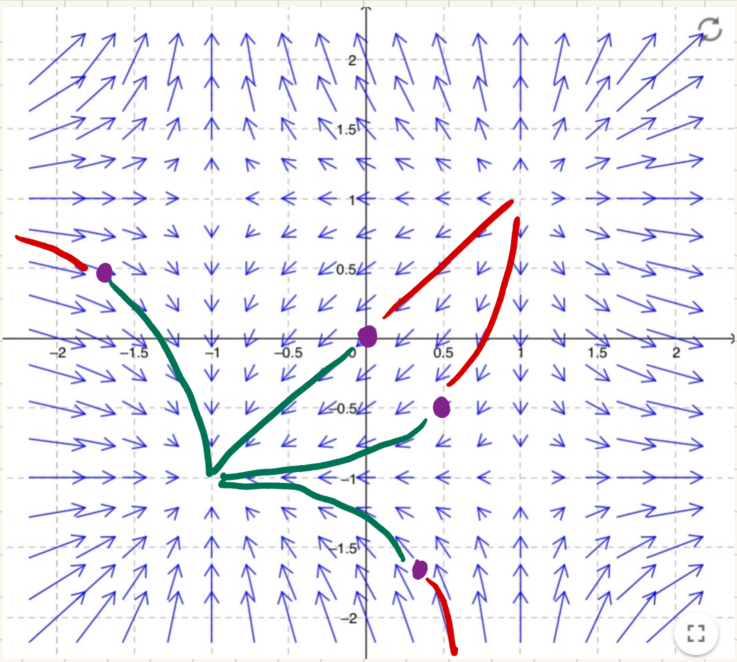
$(-1,-1)$: Local max



(At each point (a,b), I've drawn $\nabla f(a,b)$)

Here's the new perspective I want to give. Since ∇f always points in the direction of greatest increase, I should think of the gradient as pointing me in the direction I should go to find the local max or local min. If I'm at some point (a,b) in the plane, I walk in the direction the gradient points to find the local max, and walk in the opposite direction to find the local min.

For example, I'm going to draw a bunch of starting points on the gradient vector field below. will be a path that goes in the same direction as the gradient and will go against the gradient.



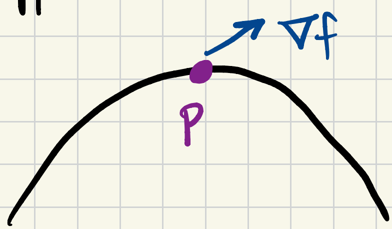
Notice that as expected, if my paths end, green paths end at my local max $(-1,-1)$, and red paths end at the local min $(1,1)$.

So, the main idea is:

$$\boxed{\text{The gradient tells me where to go to find critical points}}$$

Now let's look what happens when we add a constraint $g(x,y)=0$. Again, we'd like to use the gradient to find the critical points, but now we're trapped! Our constraint tells us that we can only move on the curve $g(x,y)=0$.

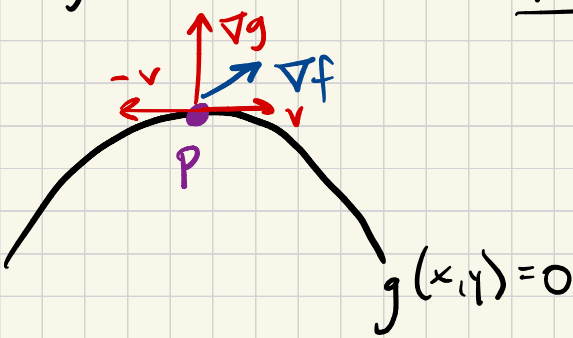
Suppose our situation looks like this:



$g(x,y)=0$

At the pt. P, the gradient tells us to move ↗ if we want to find the local max, but we can't! There are really only two directions we could move, a little to the right, or a little to the left, we <u>have</u> to stay on $g(x,y)=0$. Remember our goal, we want to maximize $f(x,y)$ while staying on $g(x,y)=0$. So since we can't

move in the direction the gradient wants us to, we need to figure out if a small step left or small step right on $g(x,y)$ will increase $f(x,y)$.
Note that these small steps can be thought of as moving in the direction tangent to $g(x,y) = 0$.



Let v denote the unit vector tangent to $g(x,y)=0$ at P, pointing right. Note that this is orthogonal to ∇g, as shown in the picture. So now our question is, if I'm at P and move in direction v or $-v$, which one increases $f(x,y)$ more? This is exactly what the directional derivative tells us!!!
So we want to find

$$D_v f(P) = \nabla f \cdot v$$

From our picture we see that $D_v f(P) > 0$ while $D_{-v} f(P) < 0$, so we need to move in the v direction!

We continue this process, making small steps along $g(x,y)=0$ that increase $f(x,y)$, i.e. moving in the direction v s.t. $D_v f(P) > 0$. So how do we know when to stop? Well, we stop when we can't move anymore, exactly when $D_v f(P) = 0$! This being 0 just means moving away from this spot will <u>not</u> increase our function.

And: $D_v f(P) = 0 \Rightarrow \nabla f(P) \cdot v = 0$

And since v is orthogonal to $\nabla g(P)$, we see that $\nabla f(P)$ and $\nabla g(P)$ point in the same direction!

So we conclude that in order to maximize or minimize $f(x,y)$ while we are stuck on $g(x,y)=0$, we just need to see where $\nabla f = \lambda \nabla g$ for some constant λ.