Proof of Lagrange Multipliers

Here we will give two arguments, one geometric and one analytic for why Lagrange multipliers work.

Critical points
For the function $w = f(x, y, z)$ constrained by $g(x, y, z) = c$ (c a constant) the critical points are defined as those points, which satisfy the constraint and where ∇f is parallel to ∇g. In equations:

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z) \quad \textit{and} \quad g(x, y, z) = c.$$

Statement of Lagrange multipliers
For the constrained system local maxima and minima (collectively extrema) occur at the critical points.

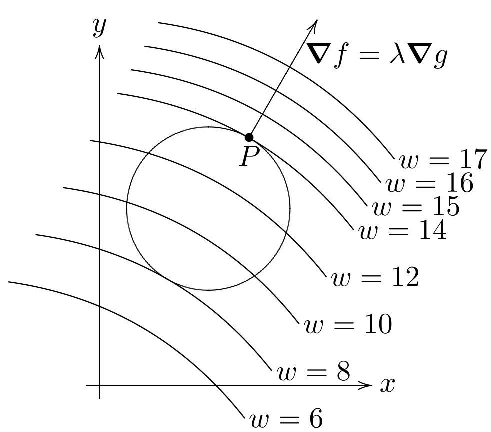
Geometric proof for Lagrange
(We only consider the two dimensional case, $w = f(x, y)$ with constraint $g(x, y) = c$.)

For concreteness, we've drawn the constraint curve, $g(x, y) = c$, as a circle and some level curves for $w = f(x, y) = c$ with explicit (made up) values. Geometrically, we are looking for the point on the circle where w takes its maximum or minimum values.

Now, start at the level curve with $w = 17$, which has no points on the circle. So, clearly, the maximum value of w on the constraint circle is less than 17. Move down the level curves until they first touch the circle when $w = 14$. Call the point where the first touch P. It is clear that P gives a local maximum for w on $g = c$, because if you move away from P in either direction on the circle you'll be on a level curve with a smaller value.

Since the circle is a level curve for g, we know ∇g is perpendicular to it. We also know ∇f is perpendicular to the level curve $w = 14$, since the curves themselves are tangent, these two gradients must be parallel.

Likewise, if you keep moving down the level curves, the last one to touch the circle will give a local minimum and the same argument will apply.



Analytic proof for Lagrange (in three dimensions)

Suppose f has a local maximum at P on the constraint surface.

Let $\mathbf{r}(t) = \langle x(t), y(t), z(t) \rangle$ be an arbitrary parametrized curve which lies on the constraint surface and has $(x(0), y(0), z(0)) = P$. Finally, let $h(t) = f(x(t), y(t), z(t))$. The setup guarantees that $h(t)$ has a maximum at $t = 0$.

Taking a derivative using the chain rule in vector form gives

$$h'(t) = \nabla f|_{\mathbf{r}(t)} \cdot \mathbf{r}'(t).$$

Since $t = 0$ is a local maximum, we have

$$h'(0) = \nabla f|_P \cdot \mathbf{r}'(0) = 0.$$

Thus, $\nabla f|_P$ is perpendicular to any curve on the constraint surface through P.

This implies $\nabla f|_P$ is perpendicular to the surface. Since $\nabla g|_P$ is also perpendicular to the surface we have proved $\nabla f|_P$ is parallel to $\nabla g|_P$. QED