• These give a tool for handling constrained max/min problems in several variables, which is actually used in applications (e.g. economics).

• Suppose you want to maximize a function \( f(x, y) \) on a curve \( g(x, y) = 0 \). One approach is to parametrize the curve by \( \vec{r}(t) = \langle x(t), y(t) \rangle \), and find stationary points of \( f(\vec{r}(t)) \): by the chain rule,

\[
0 = \frac{d}{dt} f(x(t), y(t)) = f_x(\vec{r}(t))(x'(t)) + f_y(\vec{r}(t))(y'(t)) = (\vec{\nabla} f)(\vec{r}(t)) \cdot \vec{r}'(t).
\]

If \( t = t_0 \) solves this equation, then we have \( (\vec{\nabla} f)(\vec{r}(t_0)) \perp \vec{r}'(t_0) \). But since \( \vec{\nabla} g \) is normal to level curves of \( g \), and \( \vec{r}'(t_0) \) is tangent to the level curve \( g(x, y) = 0 \) at \( \vec{r}(t_0) \), we must also have \( (\vec{\nabla} g)(\vec{r}(t_0)) \perp \vec{r}'(t_0) \). So in fact (assuming \( (\vec{\nabla} g)(\vec{r}(t_0)) \neq \vec{0} \)) \((\vec{\nabla} f)(\vec{r}(t_0))\) is parallel to \( (\vec{\nabla} g)(\vec{r}(t_0)) \), and so equals some multiple \( \lambda(\vec{\nabla} g)(\vec{r}(t_0)) \). The number \( \lambda \) is what we call the Lagrange multiplier.

• This simple argument tells us that at any local maximum (or minimum) \((x_0, y_0)\) of \( f \) on \( g(x, y) = 0 \), we have \( (\vec{\nabla} f)(x_0, y_0) = \lambda(\vec{\nabla} g)(x_0, y_0) \) for some \( \lambda \in \mathbb{R} \) (assuming \( \vec{\nabla} g \) isn’t zero there, which is essentially saying that the level curve \( g(x, y) = 0 \) isn’t “singular” there). Solving this equation (really 2 equations) together with \( g(x_0, y_0) = 0 \) therefore gives us a way to solve extremum problems without bothering to parametrize the curve \( g(x, y) = 0 \). This is great, because outside of special cases you won’t always be able to explicitly parametrize such curves, and also because it’s easier to implement on a computer.

• For 3 variables: let’s say you want to maximize or minimize \( f(x, y, z) \) subject to the constraint \( g(x, y, z) = 0 \). Then you simply set \( \vec{\nabla} f = \lambda \vec{\nabla} g \), and solve this
(together with \( g = 0 \) – altogether a system of four equations), then evaluate \( f \) at the set of “critical points” this yields.

• There are a couple general approaches to solving such systems of equations: either eliminate variables one at a time; or use 3 equations to get \( x, y, z \) in terms of \( \lambda \), substitute into the last equation to get one equation in \( \lambda \), and solve this.