Optimization

Literally, *optimization* is making something the best, but we use it in math to mean *maximization*, which is making something the biggest. (You can imagine that the thing that you're maximizing is a numerical measure of how good the thing that you're optimizing is.) Essentially the same principles apply to *minimization*, which is making something the smallest. (And *pessimization* is making something the worst, although people don't use that term very much, because who would want to do that?) A generic term for making something the largest or smallest is *extremization*.

The key principle of optimization is this:

> A quantity u can only take a maximum (or minimum) value when its differential $\mathrm{d}u$ is zero or undefined.

If you write u as $f(x, y)$, where f is a fixed differentiable function of (say) 2 variables, and x and y are quantities whose range of possible values you already understand (typically intervals), then $\mathrm{d}u = D_1 f(x, y)\,\mathrm{d}x + D_2 f(x, y)\,\mathrm{d}y$, or equivalently, $\mathrm{d}u = \frac{\partial u}{\partial x}\,\mathrm{d}x + \frac{\partial u}{\partial y}\,\mathrm{d}y$.

So one way that u might conceivably take an extreme value is if either (or both) of its partial derivatives are undefined. Another way is if both (not just one) of its partial derivatives are zero. If you can vary x and y smoothly however you please (essentially, if you are in the interior of the domain of f and you are free to access the entire domain), then these are the only possibilities. However, if you cannot vary them smoothly (essentially, if you are on the boundary of the domain of f or if the sitaution is otherwise constrained so that you cannot access the entire domain of f), then there are more possibilities!

If your constraint (or constraints) can be written as an equation $g(x, y) = 0$ (or really, with any constant on the right-hand side), then as long as the gradient ∇g is never zero on the solution set of the constraint equations, then you can use the method of *Lagrange multipliers*. Here, you set up an equation $\nabla f(x, y) = \lambda\,\nabla g(x, y)$, combine this with the equation $g(x, y) = 0$, and try to solve for x, y, and λ. (Since a vector equation is equivalent to 2 scalar equations, this amounts to a system of 3 equations in 3 variables, so there is hope to solve for it.) If you're working in 3 variables, then you might need two equations to specify the constraint, in which case there are two functions in the place of g and two Lagrange multipliers. (But you can also have just one g even in 3 dimensions; it's a question of whether the boundary in question is a surface or a curve.) While λ ultimately doesn't matter, the solutions that you get for the original variables give you additional critical points to check for extreme values.

On the other hand, you don't actually need Lagrange multipliers! Writing v for $g(x, y)$, if the constraint is $v = 0$ (or any constant), then differentiate this to get $\mathrm{d}v = 0$. (In fact, you could take any equation and just differentiate both sides.) Then if you try to solve the system of equations consisting of $\mathrm{d}u = 0$ and $\mathrm{d}v = 0$ for the differentials $\mathrm{d}x$ and $\mathrm{d}y$, you should immediately see that $\mathrm{d}x = 0$ and $\mathrm{d}y = 0$ is a solution. However, if you actually go through the steps of solving this as a system of linear equations (which you can always do because differentials are always linear in the differentials of the independent variables), you'll find that at some point you need to divide by some quantity involving x and y, which is invalid if that quantity is zero! So, setting whatever you divide by to zero and combining that with the constraint equation $v = 0$, you get two equations to solve for the two variables x and y. (With this method, λ never enters into it.) This will give you the other critical points to check for extreme values.

Be careful, because u might not have a maximum or minimum value! Assuming that u varies continuously (which it must if Calculus is to be useful at all), then it must have a maximum and minimum value whenever the domain of the function (including any constraints) is both closed and bounded (which is called *compact*); this means that if you pass continuously through the possibilities in any way, then you are always approaching some limiting possibility. However, if the range of possibilities heads off to infinity in some way, then you also have to take a limit to see what value u is approaching, which can be very difficult to do in more than one dimension. Or if there is a boundary that's not included in the domain, then you have to take a limit approaching that boundary, although in that case you can hope that you can check the boundary as if it were included, the same way as above. If any such limit is larger than every value that u actually reaches (which includes the possibility that a limit is ∞), then u has no maximum value; if any such limit is smaller than every value that u actually reaches (which includes the possibility that a limit is $-\infty$), then u has no minimum value.

So in the end, you look at these possibilities to optimize u:

- when any partial derivative of u is undefined,
- when all partial derivatives of u are zero,
- any boundary possibilities given by a constraint,
- any corners (boundaries of the boundaries) given by two constraints,
- any corners of corners given by three constraints (not possible with only two independent variables),
- etc (in more than 3 dimensions), and
- the limits approaching impossible limiting cases.

Whichever of these has the largest value of u gives you the maximum, and whichever has the smallest value of u gives you the minimum; but if the largest or smallest value is only approached in the limit, then the maximum or minimum technically does not exist.

Here is a typical problem: The hypotenuse of a right triangle (maybe it's a ladder leaning against a wall) is fixed at 20 feet, but the other two sides of the triangle could be anything. Still, since it's a right triangle, we know that $l^2 + h^2 = 20^2$, where l and h (length and height) are the lengths of legs of the triangle. (If we think of l and h as independent variables, then this equation is our constraint.) Differentiating this, $2l\,\mathrm{d}l + 2h\,\mathrm{d}h = 0$. Now suppose that we want to maximize or minimize the area of this triangle. Since it's a right triangle, the area is $A = \frac{1}{2}lh$, so $\mathrm{d}A = \frac{1}{2}h\,\mathrm{d}l + \frac{1}{2}l\,\mathrm{d}h$. If this is zero, then $\frac{1}{2}h\,\mathrm{d}l + \frac{1}{2}l\,\mathrm{d}h = 0$, to go along with the other equation $2l\,\mathrm{d}l + 2h\,\mathrm{d}h = 0$.

The equations at this point are linear in the differentials (as they always must be), so think of this is a system of linear equations in the variables $\mathrm{d}l$ and $\mathrm{d}h$. There are various methods for solving systems of linear equations; I'll use the method of addition aka elimination, but any other method should work just as well. So $\frac{1}{2}h\,\mathrm{d}l + \frac{1}{2}l\,\mathrm{d}h = 0$ becomes $2lh\,\mathrm{d}l + 2l^2\,\mathrm{d}h = 0$ (multiplying both sides by $4l$), while $2l\,\mathrm{d}l + 2h\,\mathrm{d}h = 0$ becomes $2lh\,\mathrm{d}l + 2h^2\,\mathrm{d}h = 0$ (multiplying both sides by h). Subtracting these equations gives $(2l^2 - 2h^2)\,\mathrm{d}h = 0$, so either $\mathrm{d}h = 0$ or $l^2 = h^2$. Now, l and h can change freely as long as they're positive, but we have limiting cases: $l \to 0^+$ and $h \to 0^+$. Since $l^2 + h^2 = 400$, we see that $l^2 \to 400$, so $l \to 20$, as $h \to 0$. Similarly, $h \to 20$ as $l \to 0$. In those cases, $A = \frac{1}{2}lh \to 0$. On the other hand, if $l^2 = h^2$, then $l = h$, so $l, h = 10\sqrt{2}$, since $l^2 + h^2 = 400$. In that case, $A = \frac{1}{2}lh = 100$.

So the largest area is 100 square feet, and while there is no smallest area, the area can get arbitrarily small with a limit of 0.