4 Optimization

There are two major applications of derivatives. The first, which we explored in sections 3.2 and 3.5, is to approximate functions that are hard or annoying to compute. The other is to attempt to find optimal values of functions.

The case is basically similar to the single-variable case, but as usual some extra wrinkles are introduced by having more than one input variable.

4.1 Critical points and Local Extrema

Definition 4.1. We say f has a *local maximum* at the point P_0 if $F(P_0) \geq f(P)$ for all P near P_0.

We say f has a *local minimum* at the point P_0 if $F(P_0) \leq f(P)$ for all P near P_0.

Remark 4.2. Note that we say f "has" an extremum at P. The extreme value is the actual output of f at that point. Thus, we can't say that P "is" a maximum of f.

It's possible to be very precise about what the word "near" means, but in this case we won't really bother. A point is a local maximum if you can draw a small circle around it and it gives the largest value of any point in that circle.

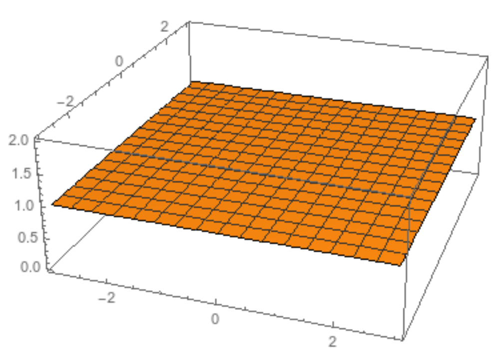
Example 4.3. Let $f(x, y) = 1$. Does this have any global maxima or minima?

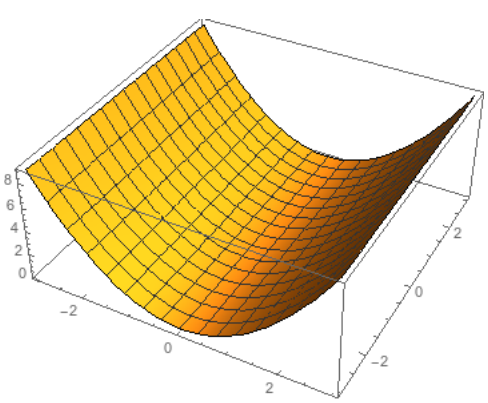
Yes. There is a maximum *and* a minimum at every single point.

This example is actually less silly in the multivariable case than in the single-variable case.

Example 4.4. Let $f(x, y) = x^2$. Does this have any global maxima or minima?

Yes. When $x = 0$ we have a local minimum whose value is zero. Thus there is a minimum at every point on this line.





With a picture (in 2 or 3 dimensions), we can identify the local extrema. And with a sufficiently simple algebraic expression we can figure out what they are. But what can we do when the situation is more complex? We need to use the derivatives.

Theorem 4.5 (Fermat)**.** *If f has a local extremum at P, and $\nabla f(P)$ exists, and P is not on the boundary of the domain of f, then $\nabla f(P) = \vec{0}$.*

Proof. Suppose $\nabla f(P) = \vec{v} \neq \vec{0}$. Then $f_{\vec{v}}(P) > 0$, so $f(P + h\vec{v}) > f(P)$ and so f doesn't have a local maximum at P. Similarly, $f(P - h\vec{v}) < f(P)$ so f doesn't have a local minimum at P. □

Definition 4.6. If $\nabla f(P) = \vec{0}$ or ∇f is undefined at P, we say that P is a *critical point* of f.

Thus Fermat's theorem tells us that all (interior) local extrema for f occur at critical points.

Counterexample where edge of domain

Example 4.7. Let $f(x, y) = -\sqrt{x^2 + y^2}$. Then

$$\nabla f(x, y) = \frac{-x}{\sqrt{x^2 + y^2}}\vec{i} + \frac{-y}{\sqrt{x^2 + y^2}}\vec{j}$$

This is actually never equal to zero, since it's undefined at the point $(0, 0)$. But this still makes the origin into a critical point, and indeed we can see that f has a local maximum at the origin.

Example 4.8. Let $f(x, y) = x^2 - 2x + y^2 - 4y + 5$

We compute

$$\nabla f(x, y) = (2x - 2)\vec{i} + (2y - 4)\vec{j}$$

which is $\vec{0}$ precisely when $(x, y) = (1, 2)$. Thus this is the only critical point.

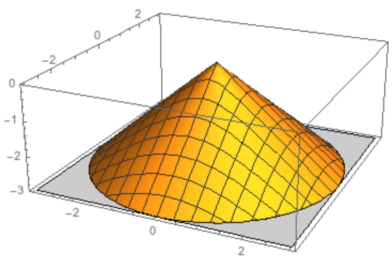
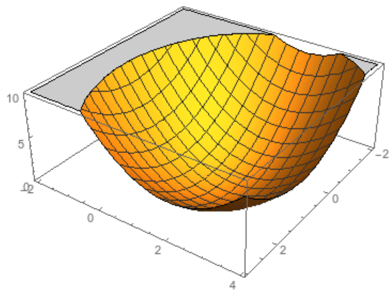
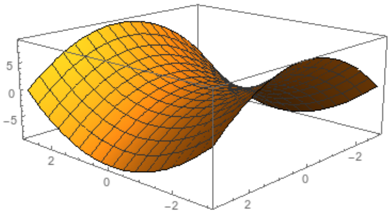
A little algebra tells us that this graph is a paraboloid $(x - 1)^2 + (y - 2)^2$. So there is a minimum at $(1, 2)$ with value 0.

Example 4.9. Let $f(x, y) = x^2 - y^2$. Then

$$\nabla f(x, y) = 2x\vec{i} - 2y\vec{j}$$

is zero when $x = y = 0$. Thus there is a single critical point at $(0, 0)$.

However, from the graph we can see that this is neither a maximum nor a minimum. In fact, it's a minimum in the x direction, and a maximum in the y direction. We call points like this "saddle points".

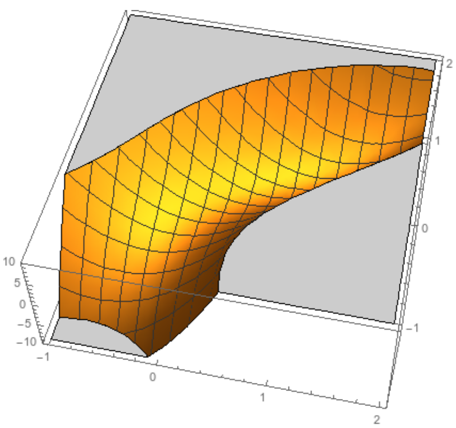
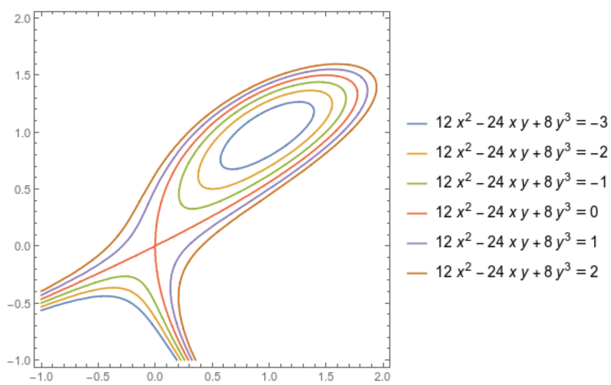




Example 4.10. Let $f(x,y) = 8y^3 + 12x^2 - 24xy$. We compute

$$\nabla f(x,y) = (24x - 24y)\vec{i} + (24y^2 - 24x)\vec{j}.$$

This is zero when $24x = 24y$ and $24y^2 = 24x$, which implies that $x = y$ and $x = y^2$, which gives us either $x = y = 0$ or $x = y = 1$. So there are two critical points, at $(0,0)$ and $(1,1)$.

From looking at the graph, we can see that there is a saddle point at $(0,0)$ and a minimum at $(1,1)$.

This last problem especially is hard to see what's happening without looking at a graph. But the second derivative can tell us what type of extrema we have at critical points.

Proposition 4.11. *Suppose* $\nabla f(a,b) = \vec{0}$. *Define*

$$D = f_{xx}(a,b)f_{yy}(a,b) - (f_{xy}(a,b))^2.$$

Then:

- *If* $D > 0$ *and* $f_{xx}(a,b) > 0$, f *has a local minimum at* (a,b).
- *If* $D > 0$ *and* $f_{xx}(a,b) > 0$, f *has a local maximum at* (a,b).
- *If* $D < 0$ f *has a saddle point at* (a,b).

Importantly, if $D = 0$ then this proposition doesn't tell us anything and we would need to do something else. We could have a local maximum, a local minimum, a saddle point, or something genuinely weird.

Example 4.12. If $f(x, y) = x^4 + y^4$, then we have

$$\begin{aligned} f_x(x,y) &= 4x^3 & f_y(x,y) &= 4y^3 \\ f_{xx}(x,y) &= 12x^2 & f_{yy}(x,y) &= 12y^2 \\ f_{xy}(x,y) &= 0 & D &= 144x^2y^2. \end{aligned}$$

We see that we have a critical point at $(0, 0)$, but at that point we get $D = 0$, which is unhelpful. But this is clearly a local minimum, since $f(0, 0) = 0$ and $f(x, y) \geq 0$.

If $f(x, y) = -x^4 - y^4$, then we have

$$\begin{aligned} f_x(x,y) &= -4x^3 & f_y(x,y) &= -4y^3 \\ f_{xx}(x,y) &= -12x^2 & f_{yy}(x,y) &= -12y^2 \\ f_{xy}(x,y) &= 0 & D &= 144x^2y^2. \end{aligned}$$

We see that we have a critical point at $(0, 0)$, but at that point we get $D = 0$, which is unhelpful. But this is clearly a local maximum, since $f(0, 0) = 0$ and $f(x, y) \leq 0$.

If $f(x, y) = x^4 - y^4$, then we have

$$\begin{aligned} f_x(x,y) &= 4x^3 & f_y(x,y) &= -4y^3 \\ f_{xx}(x,y) &= 12x^2 & f_{yy}(x,y) &= -12y^2 \\ f_{xy}(x,y) &= 0 & D &= -144x^2y^2. \end{aligned}$$

We see that we have a critical point at $(0, 0)$, but at that point we get $D = 0$, which is again unhelpful. In this case we have a saddle point: we can see that it is a minimum holding y constant, and a maximum holding x constant.

Example 4.13. Let $f(x, y) = x^2/2 + 3y^3 + 9y^2 - 3xy + 9y - 9x$.

We compute

$$\nabla(x, y) = (x - 3y - 9)\vec{i} + (9y^2 + 18y + 9 - 3x)\vec{j}$$

and thus there are critical points when $x = 3y + 9$ and $9y^2 + 18y + 9 = 3x$. Solving this gives

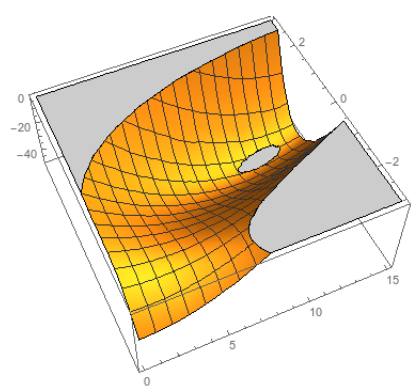
$$\begin{aligned} 9y^2 + 18y + 9 &= 9y + 27 \\ 9y^2 + 9y - 18 &= 0 \\ 9(y + 2)(y - 1) &= 0 \end{aligned}$$

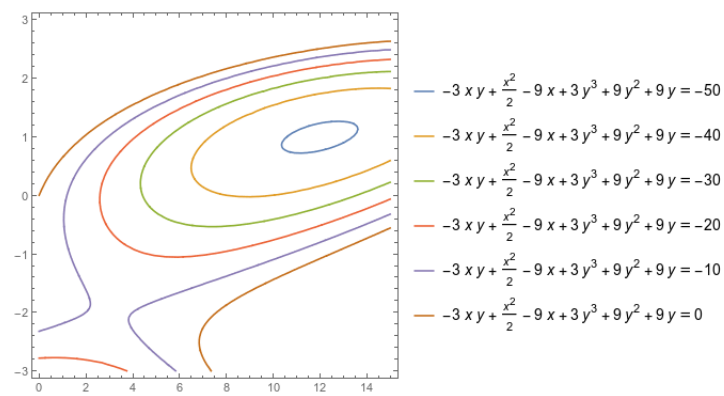
And thus $y = -2$ or $y = 1$. We see that if $y = -2$ then $x = 3$, and if $y = 1$ then $x = 12$, so the critical points are $(3, -2)$ and $(12, 1)$.

For the second derivative test, we have

$$\begin{aligned} f_{xx}(x,y) &= 1 \\ f_{yy} &= 18y + 18 \\ f_{xy}(x,y) &= -3 \\ D &= (18y+18) - (-3)^2 = 18y + 9 \\ D(3,-2) &= -27 < 0 \\ D(12,1) &= 27 \end{aligned}$$

so there is a saddle point at $(3, -2)$ and a minimum at $(12, 1)$.





4.2 Global Extrema and the Extreme Value Theorem

Critical points and the second derivative test let us determine which points are local extrema, but we often also want to know what the largest possible value we can get out of a function is.

Definition 4.14. We say f has a *global maximum* on R at the point P_0 if $F(P_0) \geq f(P)$ for all P in R.

We say f has a *global minimum* on R at the point P_0 if $F(P_0) \leq f(P)$ for all P in R.

Example 4.15. Suppose we are running a factory that produces two different products. The price we can sell each product for depends on the quantity we produce, according to the equations

$$p_1 = 600 - .3q_1 \qquad p_2 = 500 - .2q_2. \tag{1}$$

Our total cost of production is given by

$$C = 16 + 1.2q_1 + 1.5q_2 + .2q_1q_2.$$

We want to know how many of each item to produce to maximize our total profit.

Notice that here we don't really care about the relative extrema; we just want to find the best possible outcome.

First, we need to write our profit as a function of how much of each item we produce. We observe that our revenue is given by

$$R(q_1, q_2) = p_1q_1 + p_2q_2 = 600q_1 - .3q_1^2 + 500q_2 - .2q_2^2.$$

Profit is revenue minus costs, or

$$P(q_1, q_2) = R - C = -16 + 598.8q_1 - .3q_1^2 + 498.5q_2 - .2q_2^2 - .2q_1q_2.$$

Now we have P written as a function of two variables. We want to optimize it on the region $\{(q_1, q_2) : q_1 \geq 0, q_2 \geq 0\}$ since we can't produce negative quantities.

How do we find the largest possible value? In this case, the "physics" (or economics) of the situation tell us that it should occur at a relative maximum, since producing nothing is obviously suboptimal, and we expect our costs to explode as our quantity produced tends to infinity.

(Alternatively, we can notice that our equation is some sort of paraboloid and thus has a unique relative maximum that is also the absolute maximum).

Thus we look for critical points, and compute the partial derivatives.

$$\frac{\partial P}{\partial q_1} = 598.8 - .6q_1 - .2q_2$$
$$\frac{\partial P}{\partial q_2} = 498.5 - .4q_2 - .2q_1$$

and setting these equations equal to zero and solving gives us a critical point at $(q_1, q_2) = (699.1, 896.7)$. Plugging back in to equation (1) gives us prices of $(p_1, p_2) = (390.27, 320.66)$ and we get a total profit of \$432,797 dollars.

We'd like to make sure this is in fact a maximum. We can check the second partials, and we get:

$$\frac{\partial^2 P}{\partial q_1^2} = -.6 \qquad \frac{\partial^2 P}{\partial q_1 \partial q_2} = -.2 \frac{\partial^2 P}{\partial q_2^2} = -.4$$

and thus $D = (-.6)(-.4) - (-.2)^2 = .24 - .04 = .2$. Then $D > 0$ but $\frac{\partial^2 P}{\partial q_1^2} < 0$ and thus we have a local maximum. In fact, since the second derivatives are constant, we see again that we have a paraboloid; we can also infer from this that the function never increases again, so this is the only local maximum and must be a global maximum.

Example 4.16. Suppose a trucker wants to bring 480 cubic meters of gravel to a dump and needs to build a box for transport. Dumping costs \$80 per trip, plus the cost of the box.

The box has height 2m, and costs \$100 per square meter for the ends, \$50 per square meter for the sides, \$200 per square meter for the bottom. What is the optimum box size?

Let's say the box has sides of length x and ends of length y. Then the trucker takes $480/(2xy)$ trips at \$80/trip, for a total cost of $(240 \cdot 80)/(xy)$. The total cost of the box is $400y$ for the ends, $200x$ for the sides, and $200xy$ for the bottom. So total cost is

$$C = 400y + 200x + 200xy + (240 \cdot 80)/(xy) = 200\left(96/(xy) + 2y + x + xy\right).$$

We want to optimize this on the region $\{(x, y) : x > 0, y > 0\}$ since we need a positive-size box.

We can ignore the factor of 200, which doesn't change optimum. Gradient gives

$$C_x = 1 + y - 96/(x^2y) \qquad C_y = 2 + x - 96/(xy^2)$$

Setting equal to zero and solving gives

$$\begin{aligned} 96 &= x^2y + x^2y^2 & \qquad 96 &= 2xy^2 + x^2y^2 \\ x^2y &= 2xy^2 \\ x &= 2y \\ 96 &= 4y^3 + 4y^4 \end{aligned}$$

and the only positive real solution is $y = 2$. Thus the only critical point in the region is $(4, 2)$.The total cost of the transport is \$5600.

We use the second derivative test to make sure this is a minimum. (It certainly ought to be, physically). We see that

$$\begin{aligned} C_{xx} &= 192/(x^3y) & \qquad C_{xy} &= 1 + 96/(x^2y^2) & \qquad C_{yy} &= 192/(xy^3) \\ &= 6 & &= 5/2 & &= 3/2 \end{aligned}$$

and thus

$$D = 9 - 25/4 = 11/4 > 0.$$

Thus $D > 0$ and $C_{xx} > 0$, so this is a local minimum.

In both of these problems, we relied on physical intuition to tell us that a global maximum or minimum should exist. If we don't have such a clear physical setup, how can we tell?

Let's turn the question around and ask how we can *avoid* having a global maximum. One way is for the function to keep increasing infinitely the further we go in some direction. For instance, the function $f(x, y) = x + y$ doesn't have a global maximum on the plane.

Obviously this is only possible if the region is infinite. We say a region is *bounded* if it doesn't extend infinitely in any direction—that is, if we can draw a circle of finite radius around the whole region.

A more subtle way to avoid a maximum is to approach a maximum, and simply not have the point that would give you the maximum. An example here is the function $f(x, y) = x^2+y^2$ on the region $x^2 + y^2 < 1$. You can get any value less than 1, but you cannot get 1—so there is no largest possible value.

This is only possible if the region approaches but doesn't reach some point. We say a region is *closed* if it contains its entire boundary, and thus there are no points approached by the region but not contained in the region.

If a function is continuous, it turns out that these are the only way to avoid having a maximum.

Theorem 4.17 (Extreme Value)**.** *If f is a continuous function on a closed and bounded region R, then f has a global maximum and a global minimum on R.*

Thus if we have a closed and bounded region, and a continuous function, we know it must have a global maximum and a minimum.

In single variable calculus, finding these was easy. We found all the critical points and all the endpoints, plugged them into the function, and then the largest was the global maximum. In the multivariable case things are a bit harder. We still know that the global maximum must appear either at a critical point or a boundary point, but there are infinitely many boundary points so we can't just plug all of them in. Instead we need a technique to find extreme values on the boundary.

4.3 Constrained Optimization and Lagrange Multipliers

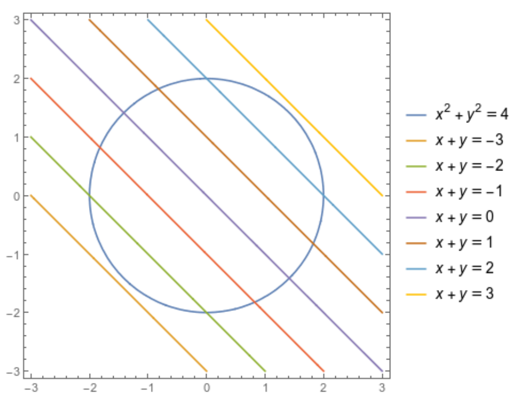
In order to answer this problem, we need to develop techniques for *constrained optimization*: optimization subject to some constraint equation. This will let us find the optimum value of a function on the boundary of its domain; it will also allow us to solve natural problems that ask us to optimize an objective function given a fixed budget, or given some physical limitations on what is possible.

Mathematically, we are looking at the problem: maximize $f(x,y)$ subject to the constraint that $g(x,y) = c$ for some constraint equation g and some constant c.

How do we do this? To find an interior maximum, we look for places where ∇f is zero. But optimizing along a constraint, we don't care if we can increase the value of f by leaving the constraint—we just need the directional derivative to be zero in the direction tangent to $g(x,y) = c$. This is equivalent to asking for ∇f to be perpendicular to that boundary.

But the boundary is just a contour or level set of g, so we know that ∇g is perpendicular to the boundary. So we're really looking for points where ∇f points in the same (or exactly opposite) direction to ∇g. We can impose this condition algebraically by looking for points where $\nabla f = \lambda \nabla g$.

Example 4.18. Let's find the maximum and minimum values of $f(x,y) = x+y$ on $x^2+y^2 = 4$.



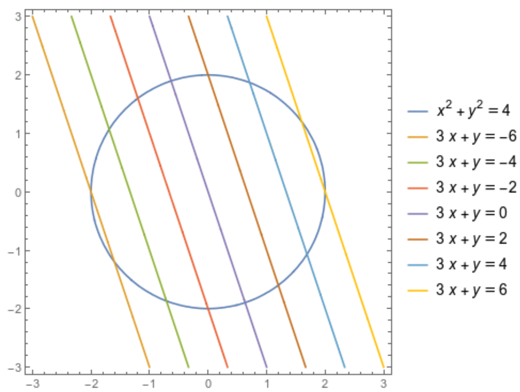
We compute $\nabla f(x,y) = (1,1)$ and $\nabla g(x,y) = (2x, 2y)$, so we are looking for points where $(1,1) = \lambda(2x, 2y)$ for some $\lambda \in \mathbb{R}$. This gives us $x = y = 1/(2\lambda)$.

To get specific values, we substitute this back into $x^2 + y^2 = 4$. We get $2x^2 = 4$ so $x^2 = 2$ and thus $x = \pm\sqrt{2}$. We know that $y = x$ so we have two critical points: $(\sqrt{2}, \sqrt{2})$ and $(-\sqrt{2}, -\sqrt{2})$. (We can also see that $\lambda = 1/(2x) = 1/\sqrt{8}$).

Plugging in values, we see that we have a maximum of $2\sqrt{2}$ at $(\sqrt{2}, \sqrt{2})$ and we have a minimum of $-2\sqrt{2}$ at $(-\sqrt{2}, -\sqrt{2})$.

Importantly, notice that this is exactly where you'd expect the maximum and minimum to be.

Example 4.19. Now let's find the maximum and minimum of $f(x,y) = 3x+y$ on $x^2+y^2 = 4$.



We compute that $\nabla f(x, y) = (3, 1)$, so we get $3 = \lambda 2x$ and $1 = \lambda 2y$. This gives us $x = 3y$, and thus we get $10y^2 = 4$, or $y = \pm\sqrt{2/5}$. So our two critical points are $(3\sqrt{2/5}, \sqrt{2/5})$ and $(-3\sqrt{2/5}, -\sqrt{2/5})$.

Plugging in values gives a maximum of $10\sqrt{2/5}$ at $(3\sqrt{2/5}, \sqrt{2/5})$ and a minimum of $-10\sqrt{2/5}$ at $(-3\sqrt{2/5}, -\sqrt{2/5})$.

We sometimes like to express these in terms of the *Lagrangian function*.

Definition 4.20. If we want to optimize $f(x, y)$ subject to $g(x, y) = c$, then the *Lagrangian function* of the problem is

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \lambda(g(x, y) - c).$$

Unconstrained critical points of $\mathcal{L}$ correspond to critical points of the original constrained optimization.

Example 4.21. Suppose we're running a factory, and our output depends on three inputs: our output is $f(x, y, z) = 20x^{3/5}y^{2/5}z^{1/5}$.

Each input has a cost. x costs 50, y costs 30, and z costs 20. If our total budget is \$18,000, how can we maximize the output?

Our budget constraint is $50x + 30y + 20z = 18000$. Then we can set up the Lagrangian:

$$\begin{aligned}
\mathcal{L}(x, y, z, \lambda) &= 20x^{3/5}y^{2/5}z^{1/5} - \lambda(50x + 30y + 20z - 18000) \\
\mathcal{L}_x &= 12\frac{y^{2/5}z^{1/5}}{x^{2/5}} - 50\lambda \\
\mathcal{L}_y &= 8\frac{x^{3/5}z^{1/5}}{y^{3/5}} + 30\lambda \\
\mathcal{L}_z &= 4\frac{x^{3/5}y^{2/5}}{z^{4/5}} + 20\lambda \\
\mathcal{L}_\lambda &= 18000 - 50x - 30y - 20z.
\end{aligned}$$

(We see that we get the constraint equation back as the partial with respect to λ). Solving for λ gives

$$\lambda = \frac{6}{25}\frac{y^{2/5}z^{1/5}}{x^{2/5}}$$
$$\lambda = \frac{4}{15}\frac{x^{3/5}z^{1/5}}{y^{3/5}}$$
$$\lambda = \frac{1}{5}\frac{x^{3/5}y^{2/5}}{z^{4/5}}$$

Solving for $z^{1/5}$ in the first two equations and setting them equal gives

$$\begin{aligned}\frac{25}{6}\lambda(x/y)^{2/5} &= \frac{15}{4}\lambda(y/x)^{3/5}\\ 50x &= 45y\\ x &= 9y/10.\end{aligned}$$

Similarly, we can solve for $x^{3/5}$ and equate the last two equations, which gives

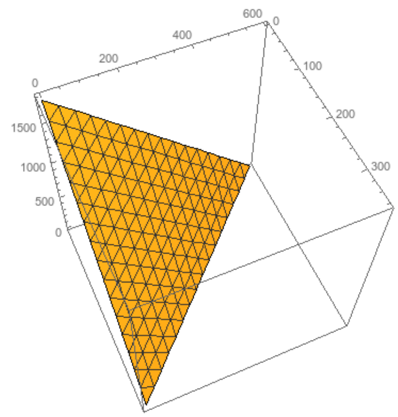
$$\begin{aligned}\frac{15}{4}\lambda y^{3/5}/z^{1/5} &= 5\lambda z^{4/5}/y^{2/5}\\ 15y &= 20z\\ z &= 3y/4.\end{aligned}$$

Plugging this all into the fourth (constraint) equation gives

$$\begin{aligned}0 &= 18000 - 50(9y/10) - 30y - 20(3y/4)\\ &= 18000 - 45y - 30y - 15y = 20000 - 90y\\ y &= 200.\end{aligned}$$

This also gives us $x = 180$ and $z = 150$.

How do we know this is a maximum? Well, the graph of the constraint is a plane, and the region of possible solutions is a triangle where the plane intersects the $x = 0$, $y = 0$, and $z = 0$ planes (since we can't produce negative amounts). On this entire boundary region the output is zero, and the output at our critical point is $\approx 462 > 0$, so we know that the boundary points are minima and the critical point is a maximum.



This brings us back to the problem of finding global extrema on a region. The basic approach is to look for critical points in the interior, and then use Lagrange multipliers to find any extrema on the boundary.

Example 4.22. Maximize and minimize $f(x, y) = (x-1)^2 + (y-2)^2$ subject to $x^2 + y^2 \leq 45$.

First we look for interior critical points. We have

$$\begin{aligned} f_x(x, y) &= 2(x-1) & x &= 1 \\ f_y(x, y) &= 2(y-2) & y &= 2 \end{aligned}$$

so the unique critical point is at $(1, 2)$.

We could use the second derivative test: we compute

$$\begin{aligned} f_{xx}(x, y) &= 2 & f_{xy} &= 0 \\ f_{yy}(x, y) &= 2 \\ D &= 4 > 0 \end{aligned}$$

and since $D > 0$, $f_{xx} > 0$ we know this is a local minimum.

But we don't actually need to do this since we're just looking for largest and smallest point. So we observe that $f(1, 2) = 0$ and move to the boundary. (It is in fact clear that this is a global minimum, since f is a sum of squares and can never give us a negative output).

On the boundary, we have the constraint $x^2 + y^2 = 45$ and we have $\nabla f(x, y) = (2(x-1), 2(y-2))$ and $\nabla g(x, y) = (2x, 2y)$. So we calculate

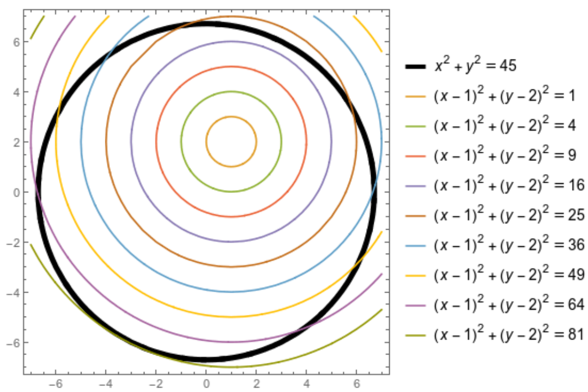
$$\begin{aligned} 2(x-1) &= \lambda 2x & \lambda &= \frac{x-1}{x} \\ 2(y-2) &= \lambda 2y & \lambda &= \frac{y-2}{y} \\ \frac{x-1}{x} &= \frac{y-2}{y} \\ xy - y &= xy - 2x & 2x &= y. \end{aligned}$$

Plugging this into the constraint gives us $5x^2 = 45$ so $x^2 = 9$ and $x = \pm 3$. Then we have $y = \pm 6$. So the two critical points are $(3, 6)$ and $(-3, -6)$.

We calculate

$$f(3, 6) = 2^2 + 4^2 = 20 \qquad f(-3, -6) = (-4)^2 + (-8)^2 = 48.$$

Thus the global maximum is 48, achieved at $(-3, -6)$, while the global minimum is 0, achieved at $(1, 2)$.



What does λ mean? It tells us how much the optimum changes when you change the constraint c. Geometrically, we have $\nabla f = \lambda \nabla g$. ∇g is, roughly speaking, how quickly c increases if we move the contour; ∇f is of course how quickly f changes when we move the contour. λ is the ratio between these, and thus how quickly f changes when we move c.

Alternatively, we can compute this with the chain rule. We know that $\frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x}$ and $\frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y}$, and of course $\frac{dg}{dc} = 1$ since $c = g(x, y)$. Then we can compute:

$$\begin{aligned} \frac{df}{dc} &= \frac{\partial f}{\partial x}\frac{dx}{dc} + \frac{\partial f}{\partial y}\frac{dy}{dc} \\ &= \lambda \frac{\partial g}{\partial x}\frac{dx}{dc} + \lambda \frac{\partial g}{\partial y}\frac{dy}{dc} = \lambda \frac{dg}{dc} = \lambda. \end{aligned}$$

Example 4.23. Let's find the global extrema of $f(x, y) = x^2 y + 3y^2 - y$ on $x^2 + y^2 \leq 10$

To find interior critical points, we compute:

$$(2xy, x^2 + 6y - 1) = (0, 0)$$

The first equation tells us that either $x = 0$ or $y = 0$. Thus the critical points are $(0, 1/6), (1, 0)$, and $(-1, 0)$. All three are in the region, so we consider all of them; we get values of $-1/12, 0,$, and 0 respectively.

Now we want to find extrema on the boundary. We compute:

$$\begin{aligned}
f_x &= 2xy = \lambda 2x = g_x \\
f_y &= x^2 + 6y - 1 = \lambda 2y = g_y \\
\lambda &= y \\
x^2 &= 2y^2 - 6y + 1 \\
10 - y^2 &= 2y^2 - 6y + 1 \\
0 &= 3y^2 - 6y - 9 = 3(y^2 - 2y - 3) = 3(y-3)(y+1)
\end{aligned}$$

So we get $y = 3$ or $y = -1$. This gives us critical points $(\pm 1, 3)$ and $(\pm 3, -1)$.

$$\begin{aligned}
f(\pm 1, 3) &= 3 + 27 - 3 = 27 \\
f(\pm 3, -1) &= 9 + 3 + 1 = 13
\end{aligned}$$

(Incidentally, we can compute that $\lambda(\pm 1, 3) = 3$, which tells us that at $(1, 3)$, increasing c by 1 would increase the maximum value of f by about 3).

So over the whole region, the global minimum is $-1/12$ at $(0, 1/6)$ and the global maximum is 27 at $(\pm 1, 3)$.

As a note: why do we always get both positive and negative x values for each y? The x variable only shows up in an x^2 so it can never affect anything whether it's positive or negative. We see this represented in the graph, because it is left-right symmetric.