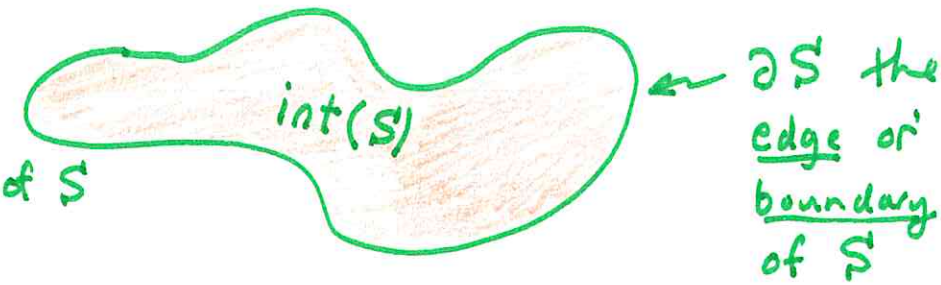
LECTURE 21 : LAGRANGE MULTIPLIERS

- based on pgs 231 - 245 of the 2020 Lecture Notes.

<u>PROBLEM</u>: Given f and some space S how can we find the min/max of the values of f attained from inputs taken from S. There are <u>two</u> cases to consider here:



int(S) = interior of S

The techniques to deal with int(S) and ∂S are different. Notice $\dim(S) = \dim(\text{int}(S))$ is <u>one</u> <u>more</u> than $\dim(\partial S)$. If S is <u>surface</u> then ∂S is <u>curve</u>. If S is a <u>volume</u> then ∂S is a <u>surface</u>. In this Lecture we share technique to study extremal values on ∂S where ∂S is either a <u>level-curve</u> or <u>level-surface</u>.

<u>Remark</u>: Looking ahead, to maximize f,

- to find extrema in int(S) find critical pts. of f where $\nabla f = 0$ (or f is discontinuous etc...) then examine Hessian to classify nature of crit. pt.
- to find extrema on ∂S either study $f(\vec{r}(t))$ where $t \longmapsto \vec{r}(t)$ parametrized ∂S. Or, use Lagrange Multipliers.

Chapter 5

optimization

The problem of optimizing a function of several variables is in many ways similar to the problem of optimization in single variable calculus. There is a fermat-type theorem; extrema are found at crtical points if anywhere. Also, there is an analogue of the closed interval method for continuous functions on some closed domain; the absolute extrema either occur at a critical point in the interior or somewhere on the boundary. However, there is no simple analogue of the first derivative test. In higher dimensions we can approach a potential extremum in infinitely many directions, in one-dimension you just have left and right approaches. The second derivative test does have a fairly simple analogue for functions of several variables. To understand the multivariate second derivative test we must first understand multivariate Taylor series. Once those are understood the second derivative test is easy to motivate. Not all instructors emphasize this point, but even in the single variable case the Taylor series expansion is probably the best tool to really understand the second derivative test. As a starting point for this chapter I assume you know what a Taylor series is, have memorized all the standard expansions and tricks, and are ready and willing to think. To the more mathematical reader, I apologize for the lack of rigor. I will not even discuss finer points of convergence or divergence. The theory of multivariate series is found in many good advanced calculus texts. I'll break from my usual format and offer the main terms in this overview:

Definition 5.0.1.

A function $f : \text{dom}(f) \subseteq \mathbb{R}^2 \to \mathbb{R}$ has a **local maximum** at (a, b) if $f(x, y) \leq f(a, b)$ for all (x, y) in some disk centered at (a, b). Likewise, $f(a, b)$ is a **local minimum** if there exists some disk D centered on (a, b) for which $(x, y) \in D$ implies $f(x, y) \geq f(a, b)$. If $S \subseteq \text{dom}(f)$ and $f(a, b) \geq f(x, y)$ for all $(x, y) \in S$ then $f(a, b)$ is a maximum of f on S. Similarly, if $S \subseteq \text{dom}(f)$ and $f(a, b) \leq f(x, y)$ for all $(x, y) \in S$ then $f(a, b)$ is a minimum of f on S. If f has a maximum or minimum on $\text{dom}(f)$ then f is said to have a **global** maximum or minimum. Maximum and minimum values are collectively called **extreme values**.

Given the terms above, let me briefly outline the chapter. In the first section we study Lagrange Multipliers which gives us a method to find extrema on constraint curves or surfaces. These constraint curves or surfaces can often be thought of as boundaries of areas or volumes. The problem of finding extrema in the interior of areas or volumes is revealed by the theory of critical points. In short, the structure of the quadratic terms in the multivariate Taylor series expanded about a critical point classify the type as maximum, minimum or saddle. However, degenerate cases such as troughs or constants require a more delicate analysis. In the second section we introduce multivariate power series and in the third section we study the second derivative test for functions of two variables. Our derivation of the second derivative test is partially based in a Lagrange

multiplier argument to circles of arbitrary radii about the critical point. See Theorem 5.1.12 and 5.1.13. This allows us to avoid some linear algebra. However, we describe in Subsection 5.1.4 how the theory of quadratic forms in linear algebra allows vast generalization of our two-dimensional result. Finally, in the fourth section we discuss the closed set test which unifies the efforts of the first two sections into a common goal.

f objective function

5.1 lagrange multipliers

The method of Lagrange Multipliers states the following: for smooth functions f, g with non-vanishing gradients[1] on $g = 0$ ← constraint eq^n.

If $f(\vec{p})$ is a maximum/minimum of f on the level-set $g = 0$ then for some constant λ

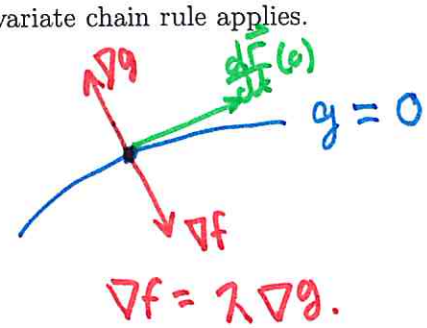
$$\boxed{\nabla f = \lambda \nabla g.}$$

Notice that the method does not provide the existence of maximums or minimums of the **objective function** f on the constraint equation $g = 0$. If no max/min for f exists on $g = 0$ then it may be possible to solve the Lagrange multiplier equation $\nabla f = \lambda \nabla g$ and find points which do not provide extrema for f on $g = 0$. We'll see examples that show that when $g = 0$ is a closed and bounded set then the extrema for f do exist. We return to this subtle points in the examples which follow the proof. Finally, I apply the method to a whole class of functions on $\mathbb{R}^2$. The last subsection is difficult but it lays the foundation for the two-dimensional second derivative test we derive later in this chapter. The logic of the test rests on a combination of the final subsection in this section and the multivariate taylor series discussed in the next section.

5.1.1 proof of the method

Proof: **($n = 2$ case)** Suppose f has a local maximum at (x_o, y_o) on the level curve $g(x, y) = 0$. Let I be an interval containing zero and $\vec{r} : I \subseteq \mathbb{R} \to \mathbb{R}^2$ be a smooth path parametrizing $g(x, y) = 0$ with $\vec{r}(0) = (x_o, y_o)$. This means $g(\vec{r}(t)) = 0$ for all $t \in I$. It is intuitively clear that the function of one-variable $h = f \circ \vec{r}$ has a maximum at $t = 0$. Therefore, by Fermat's theorem from single-variable calculus, $h'(0) = 0$. But, h is a composite function so the multivariate chain rule applies. In particular,

$$\frac{d}{dt}\Big[f(\vec{r}(t))\Big]\Big|_{t=0} = \nabla f(\vec{r}(0)) \bullet \frac{d\vec{r}}{dt}(0) = 0.$$



But, we also know $g(\vec{r}(t)) = 0$ for all $t \in I$ hence

$$\frac{d}{dt}\Big[g(\vec{r}(t))\Big] = \nabla g(\vec{r}(t)) \bullet \frac{d\vec{r}}{dt}(t) = 0.$$

for each $t \in I$. In particular, put $t = 0$ in the equation above to find

$$\frac{d}{dt}\Big[g(\vec{r}(t))\Big]\Big|_{t=0} = \nabla g(\vec{r}(0)) \bullet \frac{d\vec{r}}{dt}(0) = 0.$$

We find that both $\nabla f(x_o, y_o)$ and $\nabla g(x_o, y_o)$ are orthogonal to the tangent vector $\frac{d\vec{r}}{dt}(0)$. In two dimensions geometry forces us to conclude that $\nabla f(x_o, y_o)$ and $\nabla g(x_o, y_o)$ are colinear[2] thus there

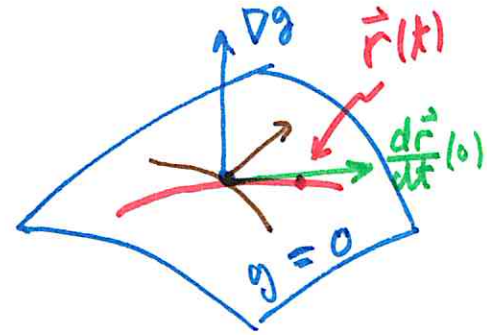
[1] this means there are no critical points for f and g on the region of interest

[2] I assume $\nabla f(x_o, y_o) \neq 0$ and $\nabla g(x_o, y_o) \neq 0$ as mentioned at the outset of this section.

exists some nonzero constant λ such that $\nabla f(x_o, y_o) = \lambda \nabla g(x_o, y_o)$. $\square$

Proof: **($n = 3$ case)** Suppose f has a local maximum at (x_o, y_o, z_o) on the level surface $g(x, y, z) = 0$. Let I be an interval containing zero and $\vec{r} : I \subseteq \mathbb{R} \to \mathbb{R}^3$ be a smooth path on the level surface $g(x, y, z) = 0$ with $\vec{r}(0) = (x_o, y_o, z_o)$. This means $g(\vec{r}(t)) = 0$ for all $t \in I$. It is intuitively clear that the function of one-variable $h = f \circ \vec{r}$ has a maximum at $t = 0$. Therefore, by Fermat's theorem from single-variable calculus, $h'(0) = 0$. But, h is a composite function so the multivariate chain rule applies. In particular,

$$\frac{d}{dt}\Big[f(\vec{r}(t))\Big]\Big|_{t=0} = \nabla f(\vec{r}(0)) \bullet \frac{d\vec{r}}{dt}(0) = 0.$$



But, we also know $g(\vec{r}(t)) = 0$ for all $t \in I$ hence

$$\frac{d}{dt}\Big[g(\vec{r}(t))\Big] = \nabla g(\vec{r}(t)) \bullet \frac{d\vec{r}}{dt}(t) = 0.$$

for each $t \in I$. In particular, put $t = 0$ in the equation above to find

$$\frac{d}{dt}\Big[g(\vec{r}(t))\Big]\Big|_{t=0} = \nabla g(\vec{r}(0)) \bullet \frac{d\vec{r}}{dt}(0) = 0.$$

$\nabla f = \lambda \nabla g.$

We find that both $\nabla f(x_o, y_o, z_o)$ and $\nabla g(x_o, y_o, z_o)$ are orthogonal to the tangent vector $\frac{d\vec{r}}{dt}(0)$. We derive this result for every smooth curve on $g(x, y, z) = 0$ thus $\nabla f(x_o, y_o, z_o)$ and $\nabla g(x_o, y_o, z_o)$ are normal to the tangent plane to $g(x, y, z) = 0$ at (x_o, y_o, z_o). It follows that $\nabla f(x_o, y_o, z_0)$ and $\nabla g(x_o, y_o, z_0)$ are colinear thus there exists some nonzero constant λ such that $\nabla f(x_o, y_o, z_0) = \lambda \nabla g(x_o, y_o, z_0)$. $\square$

In advanced calculus I discuss an more general version of the Lagrange multiplier method which solves a wider array of problems. I think these two cases suffice for calculus III. If you are curious about the general method then perhaps you should take advanced calculus.

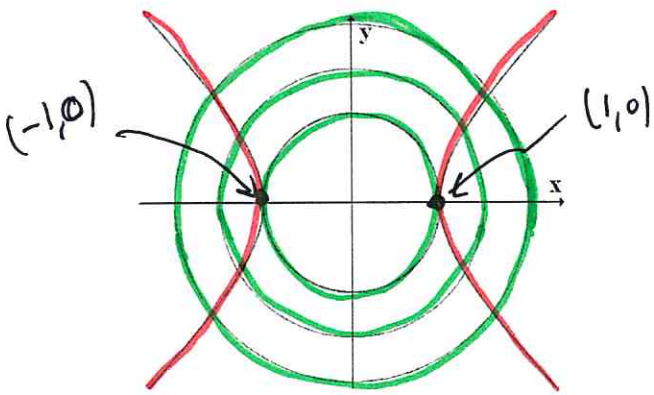
5.1.2 examples of the method

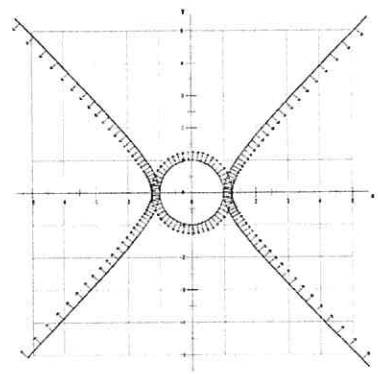
Example 5.1.1. *Suppose we wish to find maximum and minimum distance to the origin for points on the curve $x^2 - y^2 = 1$. In this case we can use the distance-squared function as our objective $f(x, y) = x^2 + y^2$ and the single constraint function is $g(x, y) = x^2 - y^2$. Observe that $\nabla f = \langle 2x, 2y \rangle$ whereas $\nabla g = \langle 2x, -2y \rangle$. We seek solutions of $\nabla f = \lambda \nabla g$ which gives us $\langle 2x, 2y \rangle = \lambda \langle 2x, -2y \rangle$. Hence $2x = 2\lambda x$ and $2y = -2\lambda y$. We must solve these equations subject to the condition $x^2 - y^2 = 1$. Observe that $x = 0$ is not a solution since $0 - y^2 = 1$ has no real solution. On the other hand, $y = 0$ does fit the constraint and $x^2 - 0 = 1$ has solutions $x = \pm 1$. Consider then*

$$2x = 2\lambda x \quad \text{and} \quad 2y = -2\lambda y \qquad \Rightarrow \qquad x(1 - \lambda) = 0 \quad \text{and} \quad y(1 + \lambda) = 0$$

Since $x \neq 0$ on the constraint curve it follows that $1 - \lambda = 0$ hence $\lambda = 1$ and we learn that $y(1 + 1) = 0$ hence $y = 0$. Consequently, $(1, 0)$ and $(-1, 0)$ are the two point where we expect to find extreme-values of f. In this case, the method of Lagrange multipliers served it's purpose, as you can see in the left graph. Below the green curves are level curves of the objective function whereas the particular red curve is the given constraint curve.

$x^2 - y^2 = 1$





The picture on the right above is a screen-shot of the Java applet created by David Lippman and Konrad Polthier to explore 2D and 3D graphs. Especially nice is the feature of adding vector fields to given objects, many other plotters require much more effort for similar visualization. See more at the website: http://dlippman.imathas.com/g1/GrapherLaunch.html. Note how the gradient vectors to the objective function and constraint function line-up nicely at those points.

In the previous example, we actually got lucky. There are examples of this sort where we could get false maxima due to the nature of the constraint function.

Example 5.1.2. *Suppose we wish to find the points on the unit circle $g(x, y) = x^2 + y^2 = 1$ which give extreme values for the objective function $f(x, y) = x^2 - y^2$. Apply the method of Lagrange multipliers and seek solutions to $\nabla f = \lambda \nabla g$:*

$$\langle 2x, -2y \rangle = \lambda \langle 2x, 2y \rangle$$

We must solve $2x = 2x\lambda$ which is better cast as $(1 - \lambda)x = 0$ and $-2y = 2\lambda y$ which is nicely written as $(1 + \lambda)y = 0$. On the basis of these equations alone we have several options:

1. *if $\lambda = 1$ then $(1 + 1)y = 0$ hence $y = 0$* $\Rightarrow x^2 + 0^2 = 1 \Rightarrow x = \pm 1 \therefore (\pm 1, 0)$
2. *if $\lambda = -1$ then $(1 - (1))x = 0$ hence $x = 0$* $\Rightarrow 0^2 + y^2 = 1 \Rightarrow y = \pm 1 \therefore (0, \pm 1)$

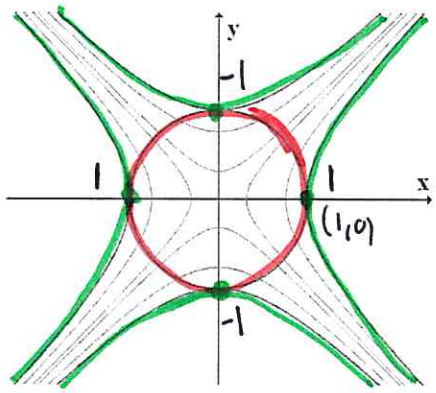
But, we also must fit the constraint $x^2 + y^2 = 1$ hence we find four solutions:

1. *if $\lambda = 1$ then $y = 0$ thus $x^2 = 1$* $\Rightarrow$ $x = \pm 1$ $\Rightarrow$ $(\pm 1, 0)$
2. *if $\lambda = -1$ then $x = 0$ thus $y^2 = 1$* $\Rightarrow$ $y = \pm 1$ $\Rightarrow$ $(0, \pm 1)$

We test the objective function at these points to ascertain which type of extrema we've located:

$$f(0, \pm 1) = 0^2 - (\pm 1)^2 = -1 \qquad \& \qquad f(\pm 1, 0) = (\pm 1)^2 - 0^2 = 1$$

When constrained to the unit circle we find the objective function attains a maximum value of 1 at the points $(1, 0)$ and $(-1, 0)$ and a minimum value of -1 at $(0, 1)$ and $(0, -1)$. Let's illustrate the answers as well as a few non-answers to get perspective. Below the green curves are level curves of the objective function whereas the particular red curve is the given constraint curve.



$g(x,y) = x^2 + y^2 = 1$

$f(x,y) = x^2 - y^2$

max is 1 at $(\pm 1, 0)$.
min is -1 at $(0, \pm 1)$.

The success of the last example was no accident. The fact that the constraint curve was a circle which is a closed and bounded subset of $\mathbb{R}^2$ means that is is a **compact** subset of $\mathbb{R}^2$. A well-known theorem of analysis states that any real-valued continuous function on a compact domain attains both maximum and minimum values. The objective function is continuous and the domain is compact hence the theorem applies and the method of Lagrange multipliers succeeds. In contrast, the constraint curve of the preceding example was a hyperbola which is not compact. We have no assurance of the existence of any extrema. Indeed, we only found minima but no maxima in Example 5.1.1.

Extreme Value Th^{m}

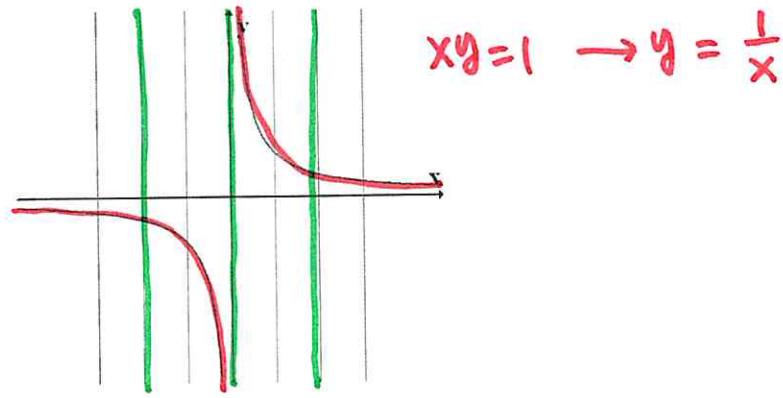
The generality of the method of Lagrange multipliers is naturally limited to smooth constraint curves and smooth objective functions. We must insist the gradient vectors exist at all points of inquiry. Otherwise, the method breaks down. If we had a constraint curve which has sharp corners then the method of Lagrange breaks down at those corners. In addition, if there are points of discontinuity in the constraint then the method need not apply. This is not terribly surprising, even in calculus I the main attack to analyze extrema of function on $\mathbb{R}$ assumed continuity, differentiability and sometimes twice differentiability. Points of discontinuity require special attention in whatever context you meet them.

At this point it is doubtless the case that some of you are, to misquote an ex-student of mine, "not-impressed". Perhaps the following examples better illustrate the dangers of non-compact constraint curves.

Example 5.1.3. *Suppose we wish to find extrema of $f(x,y) = x$ when constrained to $xy = 1$. Identify $g(x,y) = xy = 1$ and apply the method of Lagrange multipliers and seek solutions to $\nabla f = \lambda \nabla g$:*

$$\langle 1, 0 \rangle = \lambda \langle y, x \rangle \quad \Rightarrow \quad 1 = \lambda y \quad \textit{and} \quad 0 = \lambda x$$

If $\lambda = 0$ then $1 = \lambda y$ is impossible to solve hence $\lambda \neq 0$ and we find $x = 0$. But, if $x = 0$ then $xy = 1$ is not solvable. Therefore, we find no solutions. Well, I suppose we have succeeded here in a way. We just learned there is no extreme value of x on the hyperbola $xy = 1$. Below the green curves are level curves of the objective function whereas the particular red curve is the given constraint curve.



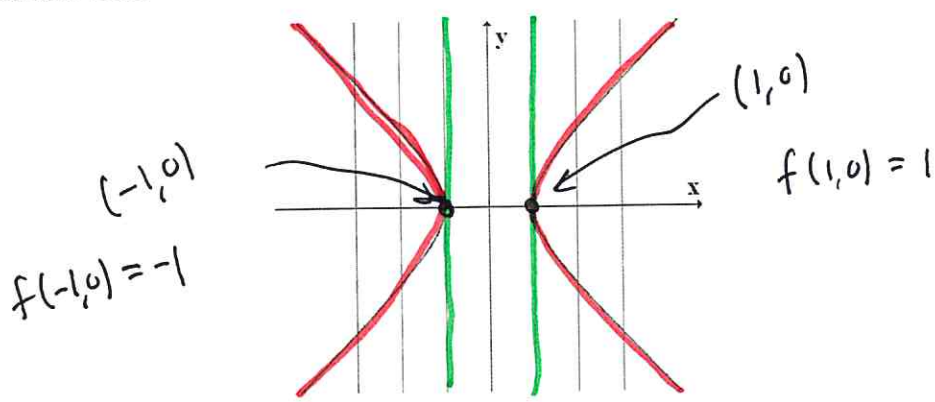
Example 5.1.4. *Suppose we wish to find extrema of $f(x,y) = x$ when constrained to $x^2 - y^2 = 1$. Identify $g(x,y) = x^2 - y^2 = 1$ and apply the method of Lagrange multipliers and seek solutions to $\nabla f = \lambda \nabla g$:*

$$\langle 1, 0\rangle = \lambda\langle 2x, -2y\rangle \quad \Rightarrow \quad 1 = 2\lambda x \quad \text{and} \quad 0 = -2\lambda y$$

If $\lambda = 0$ then $1 = 2\lambda x$ is impossible to solve hence $\lambda \neq 0$ and we find $y = 0$. If $y = 0$ and $x^2 - y^2 = 1$ then we must solve $x^2 = 1$ whence $x = \pm 1$. We are tempted to conclude that:

1. *the objective function $f(x,y) = x$ attains a maximum on $x^2 - y^2 = 1$ at $(1,0)$ since $f(1,0) = 1$*
2. *the objective function $f(x,y) = x$ attains a minimum on $x^2 - y^2 = 1$ at $(-1,0)$ since $f(1,0) = -1$*

But, both conclusions are false. Note $\sqrt{2}^2 - 1^2 = 1$ hence $(\pm\sqrt{2}, 1)$ are points on the constraint curve and $f(\sqrt{2}, 1) = \sqrt{2}$ and $f(-\sqrt{2}, 1) = -\sqrt{2}$. The error of the method of Lagrange multipliers in this context is the supposition that there exists extrema to find, in this case there are no such points. It is possible for the gradient vectors to line-up at points where there are no extrema. Below the green curves are level curves of the objective function whereas the particular red curve is the given constraint curve.



Incidentally, if you want additional discussion of Lagrange multipliers for two-dimensional problems one very nice source I certainly profitted from was the YouTube video by Edward Frenkel of Berkley. See his website http://math.berkeley.edu/ frenkel/ for links.

Example 5.1.5. *Suppose we wish to find extrema of $f(x,y) = x^2 + 3y^2$ on the unit circle $g(x,y) = x^2 + y^2 = 1$. Identify that f is the objective function and g is the contraint function for this problem. The method of Lagrange multipliers claims that extrema for f subject to $g = 1$ are found from solutions of $\nabla f = \lambda \nabla g$. In particular we face the algebra problem below:*

$$\langle 2x, 6y \rangle = \lambda \langle 2x, 2y \rangle$$

Therefore, $x = \lambda x$ and $3y = \lambda y$. We must solve simultaneously

$$x(1-\lambda) = 0, \quad y(3-\lambda) = 0, \quad x^2 + y^2 = 1$$

If $x = 0$ then $\lambda = 3$ and hence $x^2 + y^2 = 1$ implies $y = \pm 1$. On the other hand, if $\lambda = 1$ then $y = 0$ hence $x^2 + y^2 = 1$ implies $x = \pm 1$. Thus, we find the four extremal points: $(0,1), (0,-1), (1,0), (-1,0)$ and evaluation will reveal which is max/min

$$f(0, \pm 1) = 3 \qquad f(\pm 1, 0) = 1$$

Therefore, f restricted to the unit circle $x^2 + y^2 = 1$ reaches an absolute maxium value of 3 at the points $(0,-1)$ and $(0,1)$ and an absolute minimum of 1 at the points $(1,0)$ and $(-1,0)$.

I know we found the absolute maximum and minimum because the constraint curve is closes and bounded and the objective function is smooth with non-vanishing gradient near the constraint curve. These two criteria imply that extreme values exist and the method of Lagrange can find them.

Example 5.1.6. Problem: *find the closest point on the plane $\underbrace{2x - 2y + 6z}_{g(x,y,z)} = 12$ to the point $(2,3,4)$.*

Solution: *we wish to minimize the distance between the (x,y,z) on the plane and the point $(2,3,4)$. This suggests our objective function is $f(x,y,z) = (x-2)^2 + (y-3)^2 + (z-4)^2$. The constraint surface is simply $g(x,y,z) = 2x - 2y + 6z - 12 = 0$. Examine the lagrange multiplier equations:*

$$\nabla f = \lambda \nabla g \qquad \Rightarrow \qquad \langle\ 2(x-2),\ 2(y-3),\ 2(z-4)\ \rangle = \lambda \langle 2, -2, 6 \rangle$$

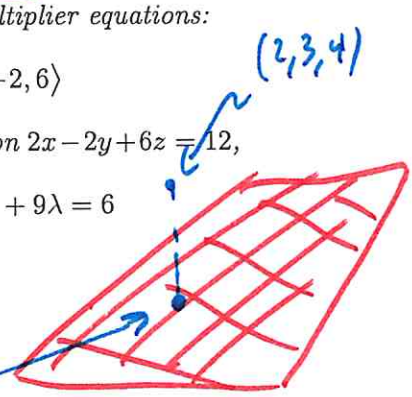
Therefore, $x = 2 - \lambda$, $y = 3 + \lambda$, $z = 4 + 3\lambda$. Substituting into the plane equation $2x - 2y + 6z = 12$,

$$2(2-\lambda) - 2(3+\lambda) + 6(4+3\lambda) = 12 \qquad \Rightarrow \qquad 2 - \lambda - 3 - \lambda + 12 + 9\lambda = 6$$

Hence, $7\lambda = 6 - 11$ so $\lambda = -5/7$. We deduce that the closest point at

$$x = 2 + \frac{5}{7} = \frac{19}{7}, \quad y = 3 - \frac{5}{7} = \frac{16}{7}, \quad z = 4 - \frac{15}{7} = \frac{13}{7}.$$

The closest point is $\boxed{(19/7,\ 16/7,\ 13/7)}$



The plane $2x - 2y + 6z = 12$ is not a bounded subset of $\mathbb{R}^3$ so we shouldn't necessarily expect to find extrema for the objective function in the last example. In fact, we found no maximally distant point. In a case such as the last example we use common sense to supplement the method. Proof of that a closest point exists involves a bit more than common sense. I'll leave it to your imagination, or a future course.

Example 5.1.7. *Let $f(x,y) = e^{xy}$ then find the extrema of f on the curve $x^3 + y^3 = 16$.*

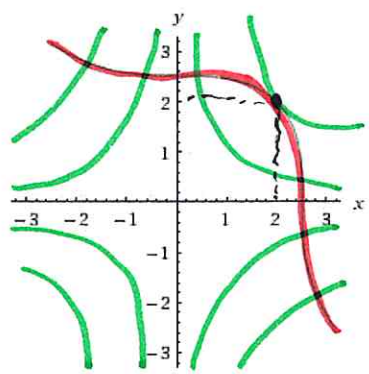
Solution: *identify our constraint as the level curve $g(x,y) = 16$ for $g(x,y) = x^3 + y^3$. The method of Lagrange multipliers suggests we solve simultaneously $g = 16$ and $\nabla f = \lambda \nabla g$. Explicitly, this yields,*

$$\langle ye^{xy},\ xe^{xy}\rangle = \lambda\langle 3x^2,\ 3y^2\rangle$$

Therefore, if we solve both component equations for λ we obtain

$$\lambda = \frac{ye^{xy}}{3x^2} = \frac{xe^{xy}}{3y^2} \quad \Rightarrow \quad \frac{y}{x^2} = \frac{x}{y^2} \quad \Rightarrow \quad y^3 = x^3.$$

Now, return to $g(x,y) = x^3 + y^3$ to see $2x^3 = 16$ hence $x = \sqrt[3]{8} = 2$. It follows that $y^3 = 8$ hence $y = 2$. Thus $f(2,2) = e^4$. This is the global maximum of $f(x,y) = e^{xy}$ on $x^3 + y^3 = 16$. This claim is seen from examining the exponential function in each quadrant as the point gets far away from the origin on the given constraint curve. The constraint curve plotted with Wolfram Alpha is given below:



$e^{xy} = c$

$xy = \ln(c)$

$y = \frac{\ln(c)}{x}$

Notice in both quadrants II. and IV. we have $xy < 0$ hence $e^{xy} < 1$. It follows the maximum found is indeed the global maximum. Also, asymptotically, the values of f approach 0 as we travel along the constraint curve far from the origin.

Example 5.1.8. *Let $f(x,y,z) = xyz$ find the extreme values of f on the surface $x^2 + 2y^2 + 3z^2 = 6$.*

Solution: *let $g(x,y,z) = x^2 + 2y^2 + 3z^2$ hence the constraint surface is given by the solution set of $g(x,y,z) = 6$. Apply the method of Lagrange multipliers to solve $g = 6$ and $\nabla f = \lambda \nabla g$ simultaneously. Explicitly,*

$$\langle yz, xz, xy\rangle = \lambda\langle 2x, 4y, 6z\rangle$$

Thus,

$$yz = 2\lambda x, \quad xz = 4\lambda y, \quad xy = 6\lambda z.$$

If any of the variables are zero then the equations above force the remaining two variables to be zero as well. Therefore, as $(0,0,0)$ is not a solution of $g(x,y,z) = 6$ we may assume $x, y, z \neq 0$ in the algebra which follows. Multiply the equations by x, y, z respectively to obtain:

$$xyz = 2\lambda x^2, \quad xyz = 4\lambda y^2, \quad xyz = 6\lambda z^2.$$

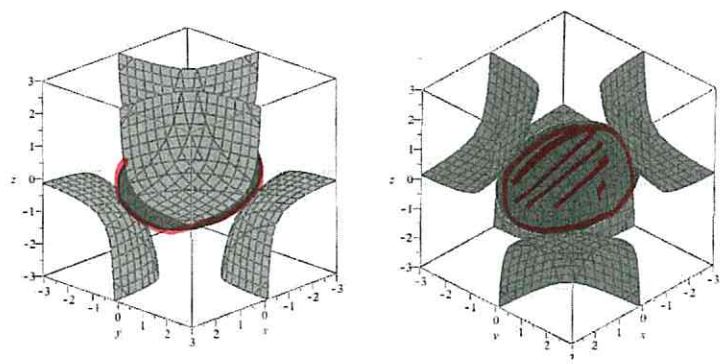
From which we find $x^2 = 2y^2 = 3z^2$ hence $g(x,y,z) = x^2 + 2y^2 + 3z^2 = 3x^2 = 6$. We find $x^2 = 2$ thus $x = \pm\sqrt{2}$ and it follows $2y^2 = 2$ hence $y = \pm 1$ and $3z^2 = 2$ thus $z = \pm\frac{\sqrt{2}}{\sqrt{3}}$. It follows we have

eight points to consider:

$$f(\sqrt{2},1,\tfrac{\sqrt{2}}{\sqrt{3}}) = \tfrac{2}{\sqrt{3}} \qquad f(\sqrt{2},1,-\tfrac{\sqrt{2}}{\sqrt{3}}) = -\tfrac{2}{\sqrt{3}}$$
$$f(\sqrt{2},-1,\tfrac{\sqrt{2}}{\sqrt{3}}) = -\tfrac{2}{\sqrt{3}} \qquad f(\sqrt{2},-1,-\tfrac{\sqrt{2}}{\sqrt{3}}) = \tfrac{2}{\sqrt{3}}$$
$$f(-\sqrt{2},1,\tfrac{\sqrt{2}}{\sqrt{3}}) = -\tfrac{2}{\sqrt{3}} \qquad f(-\sqrt{2},1,-\tfrac{\sqrt{2}}{\sqrt{3}}) = \tfrac{2}{\sqrt{3}}$$
$$f(-\sqrt{2},-1,\tfrac{\sqrt{2}}{\sqrt{3}}) = \tfrac{2}{\sqrt{3}} \qquad f(-\sqrt{2},-1,-\tfrac{\sqrt{2}}{\sqrt{3}}) = -\tfrac{2}{\sqrt{3}}.$$

The maximum value $\frac{2}{\sqrt{3}}$ *is attained on the surface at the four points which have either all positive or just two negative coordinates. The minimum value* $-\frac{2}{\sqrt{3}}$ *is attained on the surface at the four points which have an odd number of negative components. Our analysis is illustrated in diagrams below. The blue ellipsoid is the constraint and the green and red illustrate surfaces on which the objective function takes its maximum and minimum values respective:*

$f(x,y,z) = xyz$



Example 5.1.9. *Find the point on the plane* $x - y + z = 8$ *which is closest to* $(1,2,3)$.

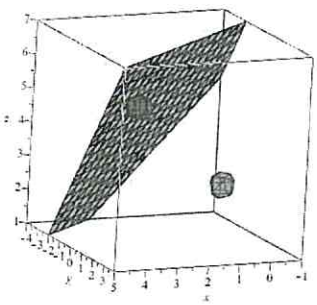
Solution: *we seek to minimize* $f(x,y,z) = (x-1)^2 + (y-2)^2 + (z-3)^2$ *subject* $g(x,y,z) = x - y + z = 2$. *Lagrange's method gives us:*

$$\nabla f = \lambda \nabla g \quad \Rightarrow \quad \begin{array}{c} 2(x-1) = \lambda \\ 2(y-2) = -\lambda \\ 2(z-3) = \lambda \end{array}$$

Thus,

$$\frac{\lambda}{2} = x - 1 = 2 - y = z - 3$$

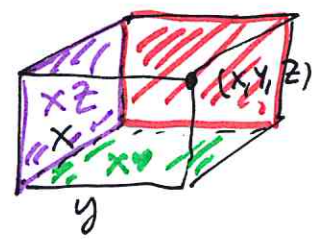
hence $x = 3 - y$ *and* $z = 5 - y$ *and we can substitute these into the plane equation to find* $3 - y - y + 5 - y = -3y + 8 = 8$ *thus* $y = 0$ *and we find* $x = 3$ *and* $z = 5$. *Therefore, the closest point on the plane is* $(3,0,5)$. *A silly picture of this is as follows: the green point is* $(1,2,3)$ *and the red is* $(3,0,5)$



Example 5.1.10. *A rectangular box without a lid is made from 12 square feet of cardboard. Find the maximum volume of such a box.*

Solution: *let x, y, z be the lengths of the sides of the box then $V = xyz$ is the volume. The contraint is given by $g(x, y, z) = 2xz + 2yz + xy = 12$. The xy is the area of the base and the top is open so this distinguishes the term from the sides of the box. We apply the method of Lagrange: consider $\nabla V = \lambda \nabla g$ yields*

$$\langle yz, xz, xy \rangle = \lambda \langle 2z + y, 2z + x, 2x + 2y \rangle.$$



Thus,

$$\begin{aligned} yz &= \lambda(2z + y) \quad \Rightarrow \quad xyz = \lambda(2zx + xy) \\ xz &= \lambda(2z + x) \quad \Rightarrow \quad xyz = \lambda(2zy + xy) \\ xy &= \lambda(2x + 2y) \quad \Rightarrow \quad xyz = \lambda(2xz + 2yz) \end{aligned}$$

Hence,

$$\lambda(2zx + xy) = \lambda(2zy + xy) = \lambda(2xz + 2yz)$$

From which we find,

$$zx = zy \qquad xy = 2xz$$

Thus $y = x$ and $y = 2z$ hence $x = 2z$ (note $x, z = 0$ are not interesting physically). The constraint equation can be reduced to an equation in z:

$$12 = 2xz + 2yz + xy = 4z^2 + 4z^2 + 4z^2 = 12z^2 \quad \Rightarrow \quad z = 1.$$

Therefore, $x = y = 2$ and we conclude the box should have dimensions $2 \times 2 \times 1$ in feet. Thus $V = 4ft^3$ is the maximum volume.

The answer above comes as no surprise, there is no difference between x, y in the problem thus by symmetry $x = y$.

Example 5.1.11. *Let $f(x,y) = xy$. Find the extrema of f on the ellipse $x^2/8 + y^2/2 = 1$.*

Solution: *identify the constraint function $g(x,y) = x^2/8 + y^2/2 = 1$. Note, $\nabla f = \lambda \nabla g$ yields*

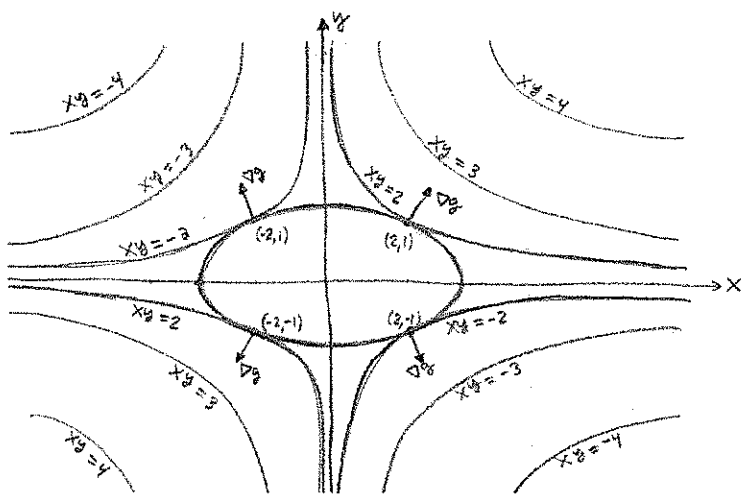
$$\langle y, x\rangle = \lambda\langle x/4, y\rangle \quad \Rightarrow \quad 4y = \lambda x \quad \& \quad x = \lambda y \quad \Rightarrow \quad y = \frac{\lambda^2 y}{4}$$

Therefore, $y(1 - \frac{\lambda^2}{4}) = 0$ from which we find solutions $y = 0$ or $\lambda = \pm 2$.

If $y = 0$ then $x = \lambda y = 0$ but $(0,0)$ is not on the ellipse.

If $\lambda = 2$ then $x = 2y$ and $4y = 2x$ and thus $x^2 = 4$ hence $x = \pm 2$ and $y = \pm 1$ hence $(-2,-1)$ and $(2,1)$ are solutions. Note $f(\pm 2, \pm 1) = (\pm 2)(\pm 1) = 2$.

If $\lambda = -2$ then $x = -2y$ and $4y = -2x$ hence $y = -\frac{1}{2}x$ thus $x^2 = 4$ and $x = \pm 2$. However, $y = -\frac{1}{2}(\pm 2) = \mp 1$. The solutions $(-2,1)$ and $(2,-1)$ follow. Note $f(\pm 2, \mp 1) = (\pm 2)(\mp 1) = -2$. In summary, we find the maximum of 2 is attained at $(-2,-1)$ and $(2,1)$ whereas the minimum of -2 is attained at $(2,-1)$ and $(-2,1)$. The picture below illustrates why:



5.1.3 extreme values of a quadratic form on a circle

In the next example we generalize the results of several past examples. In particular we intend to find the max/min for an arbitrary quadratic function in x, y on a circle of radius R. The result of this discussion will be of great use later in this chapter.

Problem: Suppose $Q(x, y) = ax^2 + 2bxy + cy^2$ for some constants a, b, c. Determine general formulas for the extrema of Q on a circle of radius R given by $g(x, y) = x^2 + y^2 = R^2$.

Solution: Apply the method of Lagrange, we seek to solve $\nabla Q = \lambda \nabla g$ subject to $g(x, y) = x^2 + y^2 = R^2$,

$$\langle 2ax + 2by, 2bx + 2cy\rangle = \lambda\langle 2x, 2y\rangle \qquad \Rightarrow \qquad ax + by = \lambda x,\ bx + cy = \lambda y$$

We must solve simultaneously the following triple of equations:

$$(a - \lambda)x + by = 0, \quad bx + (c - \lambda)y = 0, \quad x^2 + y^2 = R^2.$$

As a matrix problem, setting aside the circle equation for a moment,

$$\begin{bmatrix} a - \lambda & b \\ b & c - \lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

If $(a - \lambda)(c - \lambda) - b^2 \neq 0$ then the inverse matrix of the 2×2 exists and multiplication by of the equation above by the matrix $\frac{1}{(a-\lambda)(c-\lambda)-b^2} \begin{bmatrix} c - \lambda & -b \\ -b & a - \lambda \end{bmatrix}$ yields the solution $x = y = 0$. However, in that case there is only a solution to the circle equation $x^2 + y^2 = R^2$ if it happens that $R = 0$, we are more interested in the case $R \neq 0$ so we must look for solutions elsewhere. In other words, for interesting solutions we must insist that $(a - \lambda)(c - \lambda) - b^2 = 0$. The constants a, b, c are given so we face a quadratic equation in λ:

$$\lambda^2 - (a + c)\lambda + ac - b^2 = 0 \ \star$$

Completing the square yields solutions:

$$\lambda = \frac{a + c \pm \sqrt{(a + c)^2 - 4(ac - b^2)}}{2}$$

What type of solutions are possible from this expresssion? Simplify the expression in the radical,

$$\lambda = \frac{a + c \pm \sqrt{(a - c)^2 + 4b^2}}{2}$$

The radicand is manifestly non-negative, there is clearly no way to obtain a complex solution for real values of a, b, c. Values of λ are either zero, positive or negative. This is an important observation once we pair it with the calculation that follows. Casewise logic is needed:

Case I. Suppose $b = 0$ then $(a - \lambda)(c - \lambda) = 0$ and there are three ways to solve this:

1. $a = \lambda \neq c$ hence $y = 0$ and it follows $x = \pm R$. Therefore, extreme values of $Q(\pm R, 0) = aR^2$ are attained at $(\pm R, 0)$.

2. $c = \lambda \neq a$ hence $x = 0$ and it follows $y = \pm R$. Therefore, extreme values of $Q(0, \pm R) = cR^2$ are attained at $(0, \pm R)$.

3. $a = c = \lambda$ hence $(a - \lambda)x + by = 0$ and $bx + (c - \lambda)y = 0$ are solved, this leaves only the circle equation $x^2 + y^2 = R^2$. We find infinitely many solutions. At each point of the unit circle the value $Q(x, y) = ax^2 + cy^2 = a(x^2 + y^2) = aR^2$ is attained.

To summarize the results above, if $b = 0$ and $a \neq c$ then the extreme values of aR^2 and cR^2 are attained at the points $(\pm R, 0)$ and $(0, \pm R)$. However, if $b = 0$ and $a = c$ then Q is constant on the radius R circle with value aR^2.

Case II: Suppose $b \neq 0$. We already worked out that

$$\lambda = \frac{a + c \pm \sqrt{(a - c)^2 + 4b^2}}{2}$$

solves

$$(a - \lambda)x + by = 0, \quad bx + (c - \lambda)y = 0, \quad x^2 + y^2 = R^2.$$

Our current goal is to solve the equations above for x, y. Solve for y, I'll aim for solutions in terms of λ since we have a clear method to calculate it already,

$$y = \frac{x}{b}(\lambda - a)$$

Substitute into $x^2 + y^2 = R^2$ to find

$$x^2 + \frac{(\lambda - a)^2}{b^2}x^2 = R^2 \quad \Rightarrow \quad x_\pm = \pm R\sqrt{\frac{b^2}{b^2 + (\lambda - a)^2}}.$$

Return once more to $y = \frac{x}{b}(\lambda - a)$ to find:

$$y_\pm = \pm R\left[\frac{\lambda - a}{b}\right]\sqrt{\frac{b^2}{b^2 + (\lambda - a)^2}}.$$

Therefore, the points (x_-, y_-) and (x_+, y_+) are solutions to the Lagrange multiplier equations for each λ. Moreover, the extreme values attained via these points are given by:

$$\begin{aligned}
Q(x_\pm, y_\pm) &= ax_\pm^2 + 2bx_\pm y_\pm + cy_\pm^2 \\
&= aR^2\frac{b^2}{b^2 + (\lambda - a)^2} + 2bR^2\left[\frac{\lambda - a}{b}\right]\frac{b^2}{b^2 + (\lambda - a)^2} + cR^2\left[\frac{\lambda - a}{b}\right]^2\frac{b^2}{b^2 + (\lambda - a)^2} \\
&= \frac{R^2b^2}{b^2 + (\lambda - a)^2}\left[a + 2(\lambda - a) + c\left[\frac{\lambda - a}{b}\right]^2\right] \\
&= \frac{R^2}{b^2 + (\lambda - a)^2}\left[\lambda b^2 + b^2(\lambda - a) + c(\lambda - a)^2\right] \\
&= \frac{R^2}{b^2 + (\lambda - a)^2}\left[\lambda b^2 + (\lambda - a)\left[b^2 - ac + c\lambda\right]\right]
\end{aligned}$$

From $\star$ we know $b^2 - ac = \lambda^2 - (a+c)\lambda = \lambda^2 - a\lambda - c\lambda$ hence,

$$\begin{aligned} Q(x_\pm, y_\pm) &= \frac{R^2}{b^2 + (\lambda - a)^2}\Big[\lambda b^2 + (\lambda - a)\big[\lambda^2 - a\lambda - c\lambda + c\lambda\big]\Big] \\ &= \frac{R^2}{b^2 + (\lambda - a)^2}\Big[\lambda b^2 + \lambda(\lambda - a)^2\Big] \\ &= \lambda R^2. \end{aligned}$$

I invite the reader to prove that $b \neq 0$ implies solutions of the equation $(a-\lambda)(c-\lambda) - b^2 = 0$ are distinct. That is, given $b \neq 0$, the solutions λ_1, λ_2 must have $\lambda_1 \neq \lambda_2$. If we label these with $\lambda_1 < \lambda_2$ then it follows that $\lambda_1 R^2$ is the minimum value whereas $\lambda_2 R^2$ is the maximum value of Q on the circle $x^2 + y^2 = R^2$.

The theorem below summarizes our analysis thus far:

Theorem 5.1.12.

Suppose $Q(x,y) = ax^2 + 2bxy + cy^2$ for some constants $a, b, c \in \mathbb{R}$. Let $R > 0$ be the radius of the circle S_R with equation $x^2 + y^2 = R^2$. The **characteristic equation** $(a-\lambda)(c-\lambda) - b^2 = 0$ has only real solutions. Furthermore, the extreme values of Q on the circle are simply given by λR^2 where λ is a solution of the characteristic equation.

There are several cases implicit within the theorem above: let's denote the solutions to the characteristic equation by λ_1, λ_2,

1. if $\lambda_1 = \lambda_2$ then the Q is constant on S_R.
2. if $\lambda_1 \neq \lambda_2$ and $\lambda_1 \langle \lambda_2$ then $Q_{min} = \lambda_1 R^2$ whereas $Q_{max} = \lambda_2 R^2$.

Case (1.) is when the level curves of Q are circles. The graph $z = Q(x,y) = a(x^2 + y^2)$ either opens up ($a > 0$) or down ($a < 0$) from the origin where $Q(0,0) = 0$ is either the minimum or maxium of Q on any disk of radius R. Think geometrically for the moment, imagine shrinking $R \to 0$ to obtain this result on the disk.

Part of Case (2.) is almost the same as Case (1.) if λ_1, λ_2 share the same sign. For instance, if $0 < \lambda_1 < \lambda_2$ then $z = Q(x,y)$ opens upward with each contour being an ellipse and clearly $Q(0,0)$ is a minimum. On the other hand if $\lambda_1 < \lambda_2 < 0$ then $z = Q(x,y)$ opens downwards and each contour is an ellipse and $Q(0,0)$ is a maximum.

However, when Case (2.) has λ_1, λ_2 with different signs we find $\lambda_1 < 0 < \lambda_2$. In this case $z = Q(x,y)$ opens upward in the direction associated with λ_2 and it opens downward in the direction associated to λ_1. It has a saddle shape, and the contours of the graph are hyperbolae.

Finally, in Case (2.) if $\lambda_1 = 0$ and $\lambda_2 > 0$ then $z = Q(x,y)$ is constant along the direction corresponding to λ_1 and it opens upward along the direction corresponding to λ_2. Likwise, if $\lambda_1 = 0$ and $\lambda_2 < 0$ then $z = Q(x,y)$ is constant along the direction corresponding to λ_1 and it opens downward along the direction corresponding to λ_2.

The theorem below summarizes the relation between the characteristic equation for the quadratic form Q and its extrema in the plane $\mathbb{R}^2$. The values λ are usually called **eigenvalues** so this theorem can be essentially summarized as: the eigenvalues determine the nature of the extreme values for a quadratic form:

Theorem 5.1.13.

The graph of $z = Q(x, y) = ax^2 + 2bxy + cy^2$ for some constants $a, b, c \in \mathbb{R}$ can be categorized by real solutions of the **characteristic equation** $(a - \lambda)(c - \lambda) - b^2 = 0$. In particular,

1. if $\lambda_1, \lambda_2 > 0$ then $Q(0, 0)$ is an minimum value for Q (the graph $z = Q(x, y)$ is a paraboloid which opens up)
2. if $\lambda_1, \lambda_2 < 0$ then $Q(0, 0)$ is a maximum value for Q (the graph $z = Q(x, y)$ is a paraboloid which opens down)
3. if $\lambda_1 < 0 < \lambda_2$ then $Q(0, 0)$ is neither a maximum or minimum for Q (the graph $z = Q(x, y)$ is a hyperbolic paraboloid which opens up and down)
4. if $\lambda_1 = 0$ and $\lambda_2 > 0$ then $Q(0, 0)$ is a minimum value for Q (the graph $z = Q(x, y)$ is a parabolic trough which opens upward)
5. if $\lambda_1 = 0$ and $\lambda_2 < 0$ then $Q(0, 0)$ is a maximum value for Q (the graph $z = Q(x, y)$ is a parabolic trough which opens down)

In cases (4.) and (5.) above the local extrema is not isolated, there is a whole line on which Q is extremal. In contrast, cases (1.) and (2.) have isolated local extrema. As we apply this result later in this section cases (1-3) will play a larger role than cases (4-5).

5.1.4 quadratic forms in n-variables

optional section: I briefly explain how the last section generalizes. a good linear algebra text will provide further detail for the interested student. Notice that the last section did not use linear algebraic technique, we just brute-force solved the $n = 2$ case. To go further it is wise to learn linear algebraic techniques to organize the calculation, otherwise it could get difficult.

There is a better way to derive the results of the last section. In linear algebra we define a quadratic form on $\mathbb{R}^n$ as a function $Q : \mathbb{R}^n \to \mathbb{R}$ defined by $Q(\vec{x}) = \vec{x}^T A \vec{x}$ for a *symmetric* matrix A. It turns out that the values of Q on a sphere of radius R in $\mathbb{R}^n$ are given by the eigenvalues of A. In particular, λ is a solution to $det(A - \lambda I) = 0$ and if $\vec{x} \neq 0$ solves $(A - \lambda)\vec{x} = 0$ and $||\vec{x}|| = R$ then $Q(\vec{x}, \vec{x}) = \lambda R^2$. The value λ is called an **eigenvalue** with **eigenvector** $\vec{x}$. When you work out the details it becomes clear that $det(A - \lambda I) = 0$ is an n-th order polynomial equation in λ and, while it is not entirely trivial to prove, these solutions are all real. The list of all eigenvalues for Q is called the *spectrum*. If we order the spectrum in increasing values $\lambda_1 \leq \lambda_2 \leq \cdots \lambda_n$ then $\lambda_1 R^2$ is the minimum value whereas $\lambda_n R^2$ is the maxium value of Q on the sphere $x_1^2 + x_2^2 + \cdots + x_n^2 = R^2$ in $\mathbb{R}^n$. If you study the equations of the last section once you've studied eigenvectors and eigenvalues then you'll find that the equations provided by the Lagrange multiplier method are just the characteristic and eigenvector equations for $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$. Moreover, you learn that the big idea is that for quadratic forms is that they reduce to $Q(y_1, y_2, \ldots, y_n) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots \lambda_n y_n^2$ for appropriate choice of coordinates $(y_1, y_2, \ldots, y_n)$. For $Q(x, y) = ax^2 + 2bxy + cy^2$, the middle term $2bxy$ is an artifact of the cartesian coordinates that framed the given Q, a simple rotation will remove the non-diagonal terms in A and leave us with $Q(y_1, y_2) = \lambda_1 y_1^2 + \lambda_2 y_2^2$. Then in the (y_1, y_2) coordinates it becomes manifestly obvious that a quadratic form $Q(x, y) = ax^2 + 2bxy + cy^2$ has contours which are either hyperbolas, lines, parabolas or ellipses.