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.

Def^n/ $f : dom(f) \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$ has a local maximum at (a,b) if $f(x,y) \leq f(a,b)$ when (x,y) is on some disk centered at (a,b). Likewise $f(a,b)$ is a local minimum of f if $f(x,y) \geq f(a,b)$ when (x,y) is on some disk centered at (a,b). If we have that $f(x,y) \leq f(a,b)$ $\forall (x,y) \in S \subseteq dom(f)$ then we say that $f(a,b)$ is the maximum of f on S. Likewise if $f(x,y) \geq f(a,b)$ $\forall (x,y) \in S \subseteq dom(f)$ then $f(a,b)$ is the minimum of f on S. When $S = dom(f)$ we call those a global maximum or minimum.

5.1 lagrange multipliers

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

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

[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.

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

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

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

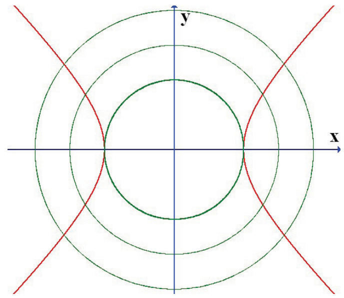
In advanced calculus I discuss an more general version of the Lagrange multiplier method which solves a wider array of problems. I think these two cases suffice for calculus III. If you are curious about the general method then perhaps you should take advanced calculus.

5.1.2 examples of the method

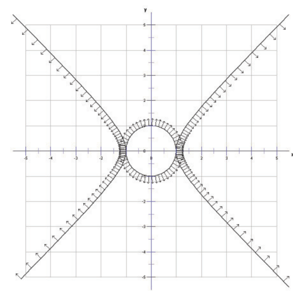
Example 5.1.1. *Suppose we wish to find maximum and minimum distance to the origin for points on the curve $x^2 - y^2 = 1$. In this case we can use the distance-squared function as our objective $f(x, y) = x^2 + y^2$ and the single constraint function is $g(x, y) = x^2 - y^2$. Observe that $\nabla f = \langle 2x, 2y \rangle$ 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*

$$2x = 2\lambda x \;\; and \;\; 2y = -2\lambda y \qquad \Rightarrow \qquad x(1 - \lambda) = 0 \;\; and \;\; y(1 + \lambda) = 0$$

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 graph. Below the green curves are level curves of the objective function whereas the particular red curve is the given constraint curve.



The picture below 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 $y = 0$*
2. *if $\lambda = -1$ then $(1-(1))x = 0$ hence $x = 0$*

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

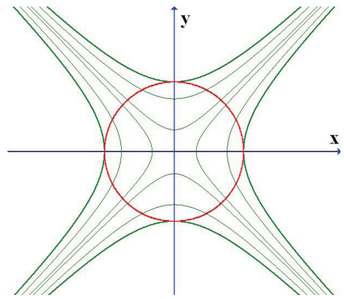
1. *if $\lambda = 1$ then $y = 0$ thus $x^2 = 1 \quad \Rightarrow \quad x = \pm 1 \quad \Rightarrow \quad (\pm 1, 0)$*

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

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.



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.

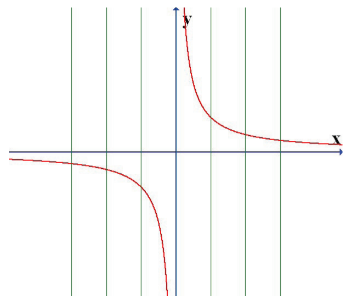
The generality of the method of Lagrange multipliers is naturally limited to smooth constraint curves and smooth objective functions. We must insist the gradient vectors exist at all points of inquiry. Otherwise, the method breaks down. If we had a constraint curve which has sharp corners then the method 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.



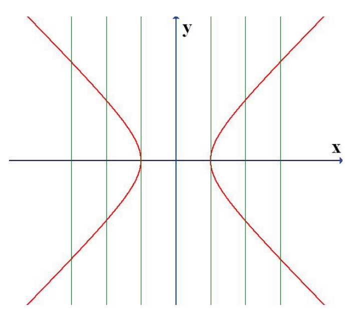
Example 5.1.4. *Suppose we wish to find extrema of $f(x, y) = x$ when constrained to $x^2 - y^2 = 1$. Identify $g(x, y) = x^2 - y^2 = 1$ and apply the method of Lagrange multipliers and seek solutions to $\nabla f = \lambda \nabla g$:*

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

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 $2x-2y+6z = 12$ 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:*

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

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. . *1*

Let $f(x,y) = e^{xy}$ then find min/max of f on the constraint surface $x^3+y^3 = 16$. Let $g(x,y) = x^3+y^3$ then the method of Lagrange multipliers indicates we solve

$$\nabla f = \lambda \nabla g \quad \text{subject to} \quad g = 16$$

Explicitly this yields,

$$\langle ye^{xy}, xe^{xy} \rangle = \lambda \langle 3x^2, 3y^2 \rangle \quad \text{with } x^3+y^3=16$$

Meaning that,

$$\left.\begin{aligned} ye^{xy} &= 3\lambda x^2 \Rightarrow \lambda = ye^{xy}/3x^2 \\ xe^{xy} &= 3\lambda y^2 \Rightarrow \lambda = xe^{xy}/3y^2 \end{aligned}\right\} \text{assuming } x,y \neq 0.$$

Can we assume $x,y \neq 0$? Well if $x=0$ then $0 = 3\lambda y^2 \Rightarrow y=0$ but then $x^3+y^3 = 0 \neq 16$. Likewise $y=0 \Rightarrow x=0 \Rightarrow x^3+y^3 \neq 16$.

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

Then $x^3+y^3 = 2x^3 = 16 \Rightarrow x^3 = 8 \therefore \underline{x=2}$ this shows that $f(2,2) = e^4$ is an extreme value of f. Notice that $x^3+y^3 = 16$ allows for $x << 0$ or $y << 0$ this will cause $\exp(xy) \longrightarrow 0$ (but not equal of course). So $\boxed{e^4 = f(2,2) \text{ is the global maximum value}}$ and there is no global minimum value (although $f(x,y)$ approaches zero asymptotically)

Example 5.1.8. . 2

$f(x,y,z) = xyz$, max/min subject to $x^2+2y^2+3z^2=6$.
that is max/min-imize w.r.t. $g(x,y,z) = x^2+2y^2+3z^2=6$.

$$\nabla f = \langle yz, xz, xy \rangle$$

$$\nabla g = \langle 2x, 4y, 6z \rangle$$

$$\nabla f = \lambda \nabla g \quad \begin{array}{l} \longrightarrow yz = 2\lambda x \longrightarrow xyz = 2\lambda x^2 \\ \longrightarrow xz = 4\lambda y \longrightarrow xyz = 4\lambda y^2 \\ \longrightarrow xy = 6\lambda z \longrightarrow xyz = 6\lambda z^2 \end{array}$$

Notice $x, y, z \neq 0$ since any one of them zero $\Rightarrow xyz = 0$ which in turn implies the other two are zero by eq^{n}'s above. Thus we can divide by $x, y, z, \ldots$

$$\frac{xyz}{\lambda} = 2x^2 = 4y^2 = 6z^2$$

$$\Rightarrow x^2 = 2y^2 = 3z^2$$

$$\Rightarrow x^2 + 2y^2 + 3z^2 = 3x^2 = 6 \quad \therefore x^2 = 2$$

$$\therefore \underline{x = \pm\sqrt{2}}$$

Now $y^2 = \frac{1}{2}x^2 = 1$ and $z^2 = \frac{1}{3}x^2 = \frac{2}{3}$ thus in total $x = \pm\sqrt{2}$, $y = \pm 1$, $z = \pm\sqrt{2}/\sqrt{3}$. The possible extreme values will be $f(\pm\sqrt{2}, \pm 1, \pm\sqrt{2}/\sqrt{3})$ by the method of Lagrange multipliers. Lets figure out which is which,

I) $f(\sqrt{2}, 1, \sqrt{2}/\sqrt{3}) = 2/\sqrt{3}$
II) $f(-\sqrt{2}, 1, \sqrt{2}/\sqrt{3}) = -2/\sqrt{3}$
III) $f(-\sqrt{2}, -1, \sqrt{2}/\sqrt{3}) = 2/\sqrt{3}$
IV) $f(-\sqrt{2}, -1, -\sqrt{2}/\sqrt{3}) = -2/\sqrt{3}$
V) $f(\sqrt{2}, -1, \sqrt{2}/\sqrt{3}) = -2/\sqrt{3}$
VI) $f(\sqrt{2}, -1, -\sqrt{2}/\sqrt{3}) = 2/\sqrt{3}$

- Maximum of $2/\sqrt{3}$ reached at cases I.), III.) and VI.)
- Min. of $-2/\sqrt{3}$ reached at cases II, IV and V.

Example 5.1.9. . *3*

Use Lagrange multipliers to find the point on $x-y+z=4$ that is closest to $(1,2,3)$.

Minimize

$$f(x,y,z) = (x-1)^2 + (y-2)^2 + (z-3)^2$$

subject to $g(x,y,z) = x-y+z = 4$. Consider then

$$\nabla f = \lambda g \longrightarrow \begin{aligned} 2(x-1) &= \lambda \\ 2(y-2) &= -\lambda \\ 2(z-3) &= \lambda \end{aligned}$$

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

$$x = 3-y$$
$$z = 5-y$$

Then $x-y+z = 3-y-y+5-y = -3y+8 = 4 \quad \therefore y = 4/3$.

thus $z = 5 - 4/3 = \frac{15-4}{3} = 11/3$ and $x = 3 - 4/3 = \frac{9-4}{3} = 5/3$

hence $(5/3, 4/3, 11/3)$ is closest point on $x-y+z=4$ to the point $(1,2,3)$.

Example 5.1.10. . 7

$\boxed{E86}$ A rectangular box without a lid is made from 12 square units of material. Find the maximum volume of such a box. That is maximize $V = xyz$ subject to $g = 2xz + 2yz + xy = 12$.

$$\nabla V = \lambda \nabla g \Rightarrow \langle yz, xz, xy \rangle = \lambda \langle 2z + y, 2z + x, 2x + 2y \rangle$$

$$yz = \lambda(2z + y) \Rightarrow xyz = \lambda(2zx + yx)$$
$$xz = \lambda(2z + x) \Rightarrow xyz = \lambda(2zy + xy)$$
$$xy = \lambda(2x + 2y) \Rightarrow xyz = \underbrace{\lambda(2xz + 2yz) = \lambda(12 - xy)}_{\text{using } g = 0}$$

Thus $\lambda(2zx + xy) = \lambda(2zy + xy) = \lambda(12 - xy)$. We can divide by λ since $\lambda = 0 \Rightarrow xyz = 0$. Note then

$$2zx + xy = 2zy + xy = 12 - xy$$
$$\Rightarrow 2zx = 2zy \Rightarrow x = y \quad \text{(note } z = 0 \text{ is not a useful value.)}$$

Next notice $\cancel{2zy} + xy = \cancel{2xz} + 2yz \Rightarrow y^2 = 2yz \therefore y = 2z$.

Then $2xz + 2yz + xy = 4z^2 + 4z^2 + 4z^2 = 12z^2 = 12$.

Hence $z = \pm 1$. Our material only comes in positive lengths so $z = 1$ hence $x = y = 2$. $\boxed{\text{The box is } 2 \times 2 \times 1.}$

Example 5.1.11. . 4

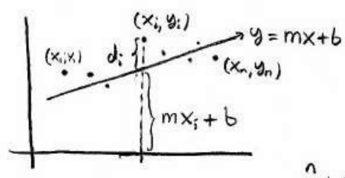
Suppose we model a relationship between x & y linearly then we expect to find m and b such that

$$y = mx + b$$

Now someone performs an experiment and collects data points

$$(x_1, y_1),\ (x_2, y_2),\ \dots,\ (x_n, y_n).$$

These points are scattered about the line, but which line? We wish to find the line that fits the data best. You've probably drawn "best fit" line in lab courses to accomplish this analysis. Recall the idea is to get an equal number of points above and below the line such that the points have the same distance in total above/below the line.



$$d_i = y_i - mx_i - b$$

Our goal is to minimize $\sum_{i=1}^{n} (d_i)^2$ with respect to m and b define $f(m,b) = \sum_{i=1}^{n} (y_i - mx_i - b)^2$. Then calculate,

$$\frac{\partial f}{\partial m} = \sum_{i=1}^{n} 2(y_i - mx_i - b)(-x_i)$$

$$\frac{\partial f}{\partial b} = \sum_{i=1}^{n} 2(y_i - mx_i - b)(-1)$$

To have $(\nabla f) = 0$ we need two things, (÷ by -2 to remove -2 factors),

$$\sum_{i=1}^{n} (x_i y_i - m x_i^2 - b x_i) = 0$$

$$\sum_{i=1}^{n} (y_i - m x_i - b) = 0$$

But this is the same as, since $\sum_{i=1}^{n} b = b \sum_{i=1}^{n} 1 = b \cdot n$,

$$\boxed{\begin{aligned} m \sum_{i=1}^{n} x_i + bn &= \sum_{i=1}^{n} y_i \\ m \sum_{i=1}^{n} x_i^2 + b \sum_{i=1}^{n} x_i &= \sum_{i=1}^{n} x_i y_i \end{aligned}}$$

I leave it to you to show these critical points do in fact give a minimum value for $f(m,b)$. (Use 2nd Der. Test)

We might consider the matrix form of our eq^{n}'s

$$A = \begin{bmatrix} \sum x_i & n \\ \sum x_i^2 & \sum x_i \end{bmatrix} \qquad d = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

Then we have $A\begin{bmatrix} m \\ b \end{bmatrix} = d$ and our sol^{n} will be $\begin{bmatrix} m \\ b \end{bmatrix} = A^{-1}d$ and since A is 2×2 we have a formula for inverse

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} \sum x_i & -n \\ -\sum x_i^2 & \sum x_i \end{bmatrix}$$

$$= \frac{1}{(\sum x_i)^2 - n\sum (x_i)^2}\begin{bmatrix} \sum x_i & -n \\ -\sum x_i^2 & \sum x_i \end{bmatrix}$$

Calculate then

$$\begin{bmatrix} m \\ b \end{bmatrix} = \frac{1}{(\sum x_i)^2 - n\sum (x_i)^2}\begin{bmatrix} \sum x_i & -n \\ -\sum x_i^2 & \sum x_i \end{bmatrix}\begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

$$= \frac{1}{(\sum x_i)^2 - n\sum (x_i)^2}\begin{bmatrix} (\sum x_i)(\sum y_i) - n\sum (x_i y_i) \\ -(\sum x_i^2)\sum (y_i) + (\sum x_i)(\sum x_i y_i) \end{bmatrix}$$

Therefore the sol^{n} is,

$$m = \frac{1}{(\sum x_i)^2 - n\sum (x_i)^2}\left[(\sum x_i)(\sum y_i) - n\sum (x_i y_i)\right]$$

$$b = \frac{1}{(\sum x_i)^2 - n\sum (x_i)^2}\left[-\sum (x_i)^2(\sum y_i) + (\sum x_i)(\sum x_i y_i)\right]$$

I suppose the graph is easier, but hey this gives <u>exactly</u> the best fit line.

Example 5.1.12. . *5*

$\boxed{\text{E84}}$ Let $f(x,y) = xy$ find extrema of f on the ellipse $\frac{x^2}{8} + \frac{y^2}{2} = 1$.
I identify that $g(x,y) = x^2/8 + y^2/2 = 1$. Consider then

$$\nabla f = \lambda \nabla g \Rightarrow \langle y, x \rangle = \lambda \langle x/4, y \rangle$$
$$\Rightarrow 4y = \lambda x \quad \text{and} \quad x = \lambda y$$
$$\Rightarrow y = \tfrac{\lambda}{4}(\lambda y)$$
$$\Rightarrow y(1 - \lambda^2/4) = 0$$
$$\Rightarrow \underline{y = 0 \quad \text{or} \quad \lambda = \pm 2}.$$

$\boxed{\text{E84 continued}}$ We've gathered that $y = 0$ or $\lambda = \pm 2$ yield the extrema,

$\underline{y=0}$ then $x = \lambda y$ and so $x = 0$ as well but $(0,0)$ not on ellipse.

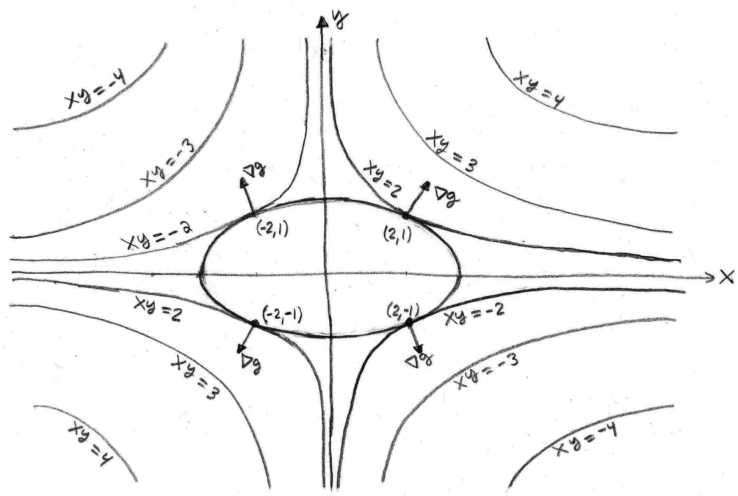
$\underline{\lambda = 2}$ $x = 2y$ and $4y = 2x$ a.k.a. $y = \frac{1}{2}x$

$$\frac{x^2}{8} + \frac{y^2}{2} = \frac{x^2}{8} + \frac{1}{2}\frac{x^2}{4} = \frac{1}{4}x^2 = 1 \quad \therefore \ x^2 = 4$$
$$\therefore \ x = \pm 2 \quad \text{and} \quad y = \tfrac{1}{2}(\pm x) = \pm 1 \quad \text{so} \quad \underline{(-2,-1) \text{ or } (2,1)}$$
$$f(-2,-1) = (-2)(-1) = 2 \quad \text{while} \quad f(2,1) = 2(1) = 2.$$

$\underline{\lambda = -2}$ $x = -2y$ and $4y = -2x$ that is $y = -\frac{1}{2}x$

$$\frac{x^2}{8} + \frac{y^2}{2} = \frac{x^2}{8} + \frac{1}{2}\frac{x^2}{4} = \frac{1}{4}x^2 = 1 \quad \therefore \ x^2 = 4$$
$$\therefore \ x = \pm 2 \quad \text{and} \quad y = \tfrac{1}{2}(\mp x) = \mp 1 \quad \text{so} \quad \underline{(-2,1) \text{ or } (2,-1)}$$
$$f(-2,1) = -2 \quad \text{while} \quad f(2,-1) = -2.$$

The extreme values are 2 and −2. The max is 2 which is reached at (−2,−1) and (2,2) while the min. is obtained at (−2,1) and (2,−1).



You can appreciate from the geometry why Lagrange's Method worked here.

Example 5.1.13. . *6*

$\boxed{E85}$ Find the point on the plane $z = x + y$ that is closest to the point $(1,1,0)$. In other words minimize $f(x,y,z) = (x-1)^2 + (y-1)^2 + z^2$ subject to $g(x,y,z) = x + y - z = 0$.

$$\nabla f = \lambda \nabla g \Rightarrow \langle 2(x-1), 2(y-1), 2z \rangle = \lambda \langle 1, 1, -1 \rangle$$

$$\Rightarrow \begin{cases} 2(x-1) = \lambda \\ 2(y-1) = \lambda \\ 2z = -\lambda \end{cases}$$

$\Rightarrow \lambda/2 = x - 1 = y - 1 = -z$

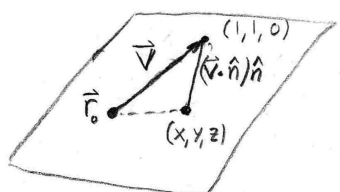
$\Rightarrow x = y$ and $z = 1 - y$

$\Rightarrow$ since $z = x + y = 2y = 1 - y$

$\therefore 3y = 1$ thus $y = 1/3 = x$ and $z = 1 - 1/3 = 2/3$.

$\boxed{\text{the closest point on the plane } z = x+y \text{ to the point } (1,1,0) \text{ is } (1/3, 1/3, 2/3)}$

<u>Lets check our answer geometrically:</u>



$\vec{V} = (1,1,0) - \vec{r}_o$

$(\vec{V} \cdot \hat{n})\hat{n} = \text{proj}_{\hat{n}}(\vec{V})$

$(x,y,z) = (1,1,0) - (\vec{V}\cdot\hat{n})\hat{n}.$

We just need to find a normal of the plane and a point on the plane. Choose $\vec{n} = \langle 1, 1, -1 \rangle$ so that $\hat{n} = \frac{1}{\sqrt{3}}\langle 1, 1, -1 \rangle$ and $\vec{r}_o = \langle 0, 0, 0 \rangle$. Hence,

$$\vec{V} = (1, 1, 0)$$

$$\text{proj}_{\hat{n}}(\vec{V}) = \frac{1}{\sqrt{3}}\langle 1,1,-1\rangle \cdot \langle 1,1,0 \rangle \hat{n} = \frac{2}{\sqrt{3}}\hat{n} = \frac{2}{3}\langle 1,1,-1 \rangle.$$

$$(1,1,0) - \text{proj}_{\hat{n}}(v) = (1,1,0) - \frac{2}{3}(1,1,-1) = \boxed{(1/3, 1/3, 2/3)}$$

this was our intuitive sol$^{\underline{n}}$ we used in the early portion of this course, the closest point falls on the normal line connecting the point and the plane.