

# Lagrange Murderpliers Done Correctly

EVAN CHEN 《陳誼廷》

8 June 2014

The aim of this handout is to provide a mathematically complete treatise on Lagrange Multipliers and how to apply them on optimization problems. The ideas here are presented logically rather than pedagogically, so it may be beneficial to read the examples before the formal statements.

Lagrange multipliers offer a way to find the extrema of an **objective function**  $f(x_1, x_2, \dots, x_n)$  subject to a **constraint function**  $g(x_1, x_2, \dots, x_n) = 0$ ; for example, we may wish to maximize  $a + b + c$  subject to  $a^2 + b^2 + c^2 = 1$ .

However, there are lots of tiny details that need to be checked in order to completely solve a problem with Lagrange multipliers. On an olympiad the use of Lagrange multipliers is almost certain to draw the wrath of graders, so it is imperative that all these details are done correctly.

## §1 Analysis Preliminaries

First, we need to define several terms. Let  $\mathbb{R}$  denote the set of real numbers, and let  $\mathbb{R}^n$  denote the set of  $n$ -tuples of real numbers; that is,  $n$ -dimensional Euclidean space.

**Definition 1.1.** A **neighborhood** of a point  $p \in \mathbb{R}^n$  is some set of points of the form

$$\{x \in \mathbb{R}^n \mid d(x, p) < \varepsilon\}$$

where  $d(x, y)$  denotes the distance between two points, and  $\varepsilon$  is a positive real number.

In other words, a neighborhood is just a “sphere” with nonzero radius  $\varepsilon$ . Now, we need to introduce some definitions from real analysis.

**Definition 1.2.** A subset  $S$  of  $\mathbb{R}^n$  is **open** in  $\mathbb{R}^n$  if for every point  $p \in S$ , there exists a neighborhood of  $p$  completely contained inside  $S$ .

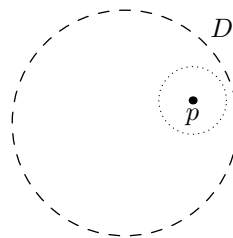


Figure 1: The open disk  $D$  is open.

**Example 1.3**

In  $\mathbb{R}^2$ , the open disk

$$D = \{(x, y) \mid x^2 + y^2 < 1\}$$

is open because for every point  $p$  inside it, we can find a small neighborhood completely inside the disk. See [Figure 1](#), noting that  $D$  does not contain its “boundary”.

Similarly, the open interval  $(0, 1)$  is open in  $\mathbb{R}$ .

**Definition 1.4.** A subset  $S$  of  $\mathbb{R}^n$  is **closed** in  $\mathbb{R}^n$  if the following holds: if  $x_1, x_2, \dots$ , is a sequence of points in  $S$  which converges to some limit  $x$ , then  $x \in S$ .

An equivalent definition is that  $\mathbb{R}^n \setminus S$  should be open.

**Example 1.5**

The closed disk  $\overline{D} = \{(x, y) \mid x^2 + y^2 \leq 1\}$  is closed. Similarly, the closed interval  $[0, 1]$  is closed in  $\mathbb{R}$  (meaning: if a convergent sequence of reals is between 0 and 1 inclusive, then it must converge to a real number between 0 and 1 inclusive).

And now you know why open and closed intervals are so called! Note that most sets are neither open nor closed in  $\mathbb{R}^n$ , and some sets are both open and closed (such as  $\emptyset$ , the empty set, as well as the entire  $\mathbb{R}^n$ ).

**Remark.** Up to now we have been kōsher in saying that a set  $S$  is closed **in**  $\mathbb{R}^n$ ; this is important because a set is only open or closed relative to its parent. However, we will henceforth drop the clause “in  $\mathbb{R}^n$ ” for brevity.

To aid with constructing closed and open sets, we have the following nice statements.

**Proposition 1.6**

Let  $S$  and  $T$  be sets in  $\mathbb{R}^n$ . If  $S$  and  $T$  are both open (resp. closed), then  $S \cup T$  and  $S \cap T$  are also both open (resp. closed).

**Proposition 1.7**

Let  $S \subseteq \mathbb{R}^m$  and  $T \subseteq \mathbb{R}^n$ . If  $S$  and  $T$  are both open (resp. closed), then  $S \times T$  is open (resp. closed) in  $\mathbb{R}^{m+n}$ .

You may also have noticed by now that open and closed sets seem to differ by a “boundary”. We can formalize this with the notion of a closure.

**Definition 1.8.** Given a set  $U$  the **closure**  $\overline{U}$  is the smallest closed set containing  $U$ .

I’ve cheated here with the notion of “smallest”; one can prove that “smallest by inclusion” works out. In practice, for  $\mathbb{R}^n$  this just means e.g. adding in the boundary points back into an open set.

Finally, we need one other notion, called compactness.

**Definition 1.9.** A subset  $K$  of  $\mathbb{R}^n$  is **compact** if it is closed and *bounded*, i.e. can be

contained inside a rectangular prism with finite dimensions.<sup>1</sup>

### Example 1.10

The rectangular box  $B = [0, 1] \times [0, 2] \times [0, 3]$  in  $\mathbb{R}^3$  is compact.

## §2 Lagrange Multipliers

We can give the statement of the theorem of Lagrange Multipliers.

### Theorem 2.1 (Lagrange Multipliers)

Let  $U$  be an open subset of  $\mathbb{R}^n$ , and let  $f: U \rightarrow \mathbb{R}$  and  $g: U \rightarrow \mathbb{R}$  be continuous functions with continuous first derivatives. Define the constraint set

$$S = \{\mathbf{x} \in U \mid g(\mathbf{x}) = c\}$$

for some real number  $c$ .

Suppose  $\mathbf{x} \in S$  is a **local extrema** of  $f$  along  $S$ , meaning  $\mathbf{x}$  gives minimal or maximal value of  $f$  among points in  $S$  for *some* neighborhood around  $\mathbf{x}$ . Then either  $\nabla g(\mathbf{x}) = \mathbf{0}$  at this point, or for some real number  $\lambda$ ,

$$\nabla f(\mathbf{x}) = \lambda \nabla g(\mathbf{x}).$$

This is fine and well, but it does not tell us anything about the **global maximum** of  $f$  constrained to  $g$ . Indeed, such a global maximum may not even exist! That means there is still tons of work left to find the maximum value of  $f$ , if indeed it exists.

The following two lemmas let us save face.

### Lemma 2.2 (Compact sets have unique maximums)

Let  $K$  be a compact set and  $f: K \rightarrow \mathbb{R}$  be a continuous function. Then  $f$  has a **global** maximum over  $K$  – there exists a point  $\mathbf{x} \in K$  such that  $f(\mathbf{x}) \geq f(\mathbf{y})$  for any other point  $\mathbf{y} \in K$ .

### Lemma 2.3 (Pre-images are closed sets)

If  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  is a continuous function, then the set  $C = \{\mathbf{x} \in \mathbb{R}^n \mid g(\mathbf{x}) = c\}$  is a closed set for any real number  $c$ .

Hence, we can escape using the following method. Let's take our function  $f: U \rightarrow \mathbb{R}$  with  $U$  open, and consider the closure  $\bar{U}$  of  $U$ . Let's instead optimize  $f: \bar{U} \rightarrow \mathbb{R}^n$ , where  $\bar{U}$  is the closure of  $U$ . Define

$$C = \{\mathbf{x} \in \mathbb{R}^n \mid g(\mathbf{x}) = c\}.$$

Note that  $C$  is closed by our lemma. Also, define

$$S = U \cap C = \{\mathbf{x} \in U \mid g(\mathbf{x}) = c\}$$

<sup>1</sup>This is not the “real” definition of compactness, it's the one that holds in  $\mathbb{R}^n$ . The actual definition in weirder spaces  $X$  is that any cover of  $X$  with open sets has a finite subcover.

as well as

$$\overline{S} = \overline{U} \cap C = \{\mathbf{x} \in \overline{U} \mid g(\mathbf{x}) = c\}.$$

As the notation suggests,  $\overline{S}$  is the closure of  $S$ , and in any case is clearly itself closed. If either of  $C$  or  $\overline{U}$  is bounded, then so is  $\overline{S}$ , meaning  $\overline{S}$  is compact! When that occurs,  $f$  achieves a maximum  $\mathbf{x} \in \overline{S}$  somewhere on the set  $C \cap \overline{U}$ .

We can then consider two cases.

- If  $\mathbf{x} \in \overline{U} - U$ , that is,  $\mathbf{x}$  lies in the “boundary”, then we can check this case manually.
- If  $\mathbf{x} \in U$ , that is,  $\mathbf{x}$  lies in the “interior”, then we can use Lagrange multipliers. After all, if  $\mathbf{x}$  is a global maximum, then it is certainly a local one as well. Here the constraint set is simply  $S = C \cap U$ .

This is the method of Lagrange multipliers.

## §3 Examples

### §3.1 An Easy Example

#### Example 3.1 (AM-GM)

For positive reals  $a, b, c$ , prove that

$$a + b + c \geq 3\sqrt[3]{abc}.$$

*Solution.* The first step is de-homogenizing the inequality to assume that  $a + b + c = 3$ , in which case we wish to prove  $abc \leq 1$ . Thus, define

$$f(a, b, c) = abc$$

and

$$g(a, b, c) = a + b + c.$$

Note that  $\nabla g = \langle 1, 1, 1 \rangle \neq \mathbf{0}$  at all points, and moreover that  $f$  and  $g$  are continuous with continuous partial derivatives.

We will let  $U = (0, 3)^3$ ,  $\overline{U} = [0, 3]^3$  (meaning we will actually prove the inequality for all  $a, b, c \geq 0$ ). Note that  $\overline{U}$  is bounded, so the set

$$\overline{S} = \{\mathbf{x} \in \overline{U} : g(\mathbf{x}) = 3\}$$

is closed and hence compact. So  $f$  achieves a maximum value on  $\overline{S}$ , say  $\mathbf{x}$ .

If  $\mathbf{x}$  lies on the boundary, then at least one component of  $\mathbf{x}$  is zero, whence  $abc = 0$ ; this is clearly not a maximum.

Otherwise, let  $\mathbf{x} = (a, b, c)$  be a point of  $S = \{\mathbf{x} \in U : g(\mathbf{x}) = 3\}$ . By Lagrange multipliers, we know

$$\nabla f(a, b, c) = \lambda \nabla g(a, b, c)$$

or

$$\langle bc, ca, ab \rangle = \lambda \langle 1, 1, 1 \rangle.$$

That is,  $bc = ca = ab = \lambda$ . If  $\lambda = 0$ , then  $a = b = c = 0$ , and  $f$  is zero at such points. Otherwise, we conclude that  $a = b = c$ . Since  $a + b + c = 3$ , we must have  $a = b = c = 1$ . Hence  $f(1, 1, 1) = 1$ .

We find that the critical point yielding the largest value on  $\overline{S}$  is  $(1, 1, 1)$ , which gives  $f(1, 1, 1) = 1$ . This must be a maximum, and hence  $f(a, b, c) \leq 1$  for all  $a + b + c = 3$ ,  $0 \leq a, b, c \leq 3$ , as desired.  $\square$

### §3.2 A Hard Example

Now we try a much bloodier example.

#### Example 3.2 (USAMO 2001/3)

Let  $a, b, c$  be nonnegative reals with  $a^2 + b^2 + c^2 + abc = 4$ . Prove that

$$0 \leq ab + bc + ca - abc \leq 2.$$

*Solution.* The left-hand side of the inequality is trivial; just note that  $\min\{a, b, c\} \leq 1$ . Hence, we focus on the right side. We use Lagrange Multipliers.

Define

$$U = \{(a, b, c) \mid a, b, c > 0 \text{ and } a^2 + b^2 + c^2 < 1000\}.$$

This is an intersection of open sets, so it is open. Its closure is

$$\bar{U} = \{(a, b, c) \mid a, b, c \geq 0 \text{ and } a^2 + b^2 + c^2 \leq 1000\}.$$

Hence the constraint set

$$\bar{S} = \{\mathbf{x} \in \bar{U} : g(\bar{\mathbf{x}}) = 4\}$$

is compact, where  $g(a, b, c) = a^2 + b^2 + c^2 + abc$ .

Define

$$f(a, b, c) = a^2 + b^2 + c^2 + ab + bc + ca.$$

It's equivalent to show that  $f \leq 6$  subject to  $g$ . Over  $\bar{S}$ , it must achieve a global maximum. Now we consider two cases.

If  $\mathbf{x}$  lies on the boundary, that means one of the components is zero (since  $a^2 + b^2 + c^2 = 1000$  is clearly impossible). WLOG  $c = 0$ , then we wish to show  $a^2 + b^2 + ab \leq 6$  for  $a^2 + b^2 = 4$ , which is trivial.

Now for the interior  $U$ , we may use the method of Lagrange Multipliers. Consider a local maximum  $\mathbf{x} \in U$ . Compute

$$\nabla f = \langle 2a + b + c, 2b + c + a, 2c + a + b \rangle$$

and

$$\nabla g = \langle 2a + bc, 2b + ca, 2c + ab \rangle.$$

Of course,  $\nabla g \neq \mathbf{0}$  everywhere, so introducing our multiplier yields

$$\langle 2a + b + c, a + 2b + c, a + b + 2c \rangle = \lambda \langle 2a + bc, 2b + ca, 2c + ab \rangle.$$

Note that  $\lambda \neq 0$  since  $a, b, c > 0$ . Subtracting  $2a + b + c = \lambda(2a + bc)$  from  $a + 2b + c = \lambda(2b + ca)$  implies that

$$(a - b)([2\lambda - 1] - \lambda c) = 0.$$

We can derive similar equations for the others. Hence, we have three cases.

1. If  $a = b = c$ , then  $a = b = c = 1$ , and this satisfies  $f(1, 1, 1) \leq 6$ .
2. If  $a, b, c$  are pairwise distinct, then we derive  $a = b = c = 2 - \lambda^{-1}$ , contradiction.

3. Now suppose that  $a = b \neq c$ . That means  $\lambda = \frac{1}{2-a}$  (of course our conditions force  $c < 2$ ). Now

$$2a + 2c = a + b + 2c = \lambda(2c + ab) = \frac{1}{2-a}(2c + a^2)$$

which implies

$$4a + 4c - 2a^2 - 2ac = 2c + a^2$$

meaning (with the additional note that  $a \neq 1$ ) we have

$$c = \frac{3a^2 - 4a}{2 - 2a}.$$

Note that at this point,  $c > 0$  forces  $1 < a < \frac{4}{3}$ .

The constraint  $a^2 + b^2 + c^2 + abc = 4 \iff c^2 + a^2c + (2a^2 - 4) = 0$  now gives

$$(3a^2 - 4a)^2 + a^2(3a^2 - 4a)(2 - 2a) + (2a^2 - 4)(2 - 2a)^2 = 0.$$

Before expanding this, it is prudent to see if it has any rational roots. A quick inspection finds that  $a = 2$  is such a root (precisely,  $16 - 32 + 16 = 0$ ). Now, we can expand and try to factor:

$$\begin{aligned} 0 &= -6a^5 + 31a^4 - 48a^3 + 8a^2 + 32a - 16 \\ &= (a - 2)(-6a^4 + 19a^3 - 10a^2 - 12a + 8) \\ &= (a - 2)^2(-6a^3 + 7a^2 + 4a - 4) \\ &= (a - 2)^2(2 - 3a)(2a^2 - a - 2). \end{aligned}$$

The only root  $a$  in the interval  $(1, \frac{4}{3})$  is  $a = \frac{1}{4}(1 + \sqrt{17})$ . To finish, write

$$c = \frac{a(3a - 4)}{2 - 2a} = \frac{1}{8}(7 - \sqrt{17})$$

and

$$f(a, b, c) = 3a^2 + 2ac + c^2 = \frac{1}{32}(121 + 17\sqrt{17}).$$

This is the last critical point, so we're done once we check this is less than 6. This follows from the inequality  $17^3 < (6 \cdot 32 - 121)^2$ ; in fact, we actually have

$$\frac{1}{32}(121 + 17\sqrt{17}) \approx 5.97165.$$

This completes the solution. □

### §3.3 A Non-Example

The following is not actually an example because you get stuck at the end (and it turns out you *must* get stuck, as I explain), but I included it because of popular request.

**Example 3.3** (USAMO 2014/1)

Let  $w, x, y, z$  be real numbers satisfying

$$wx + xy + yz + zw + wy + xz \geq 5 + wxyz.$$

Prove that

$$(w^2 + 1)(x^2 + 1)(y^2 + 1)(z^2 + 1) \geq 16.$$

*Non-Solution.* By scaling appropriately, we can assume our condition is actually an equality. Moreover, we will only prove the inequality in the case where

$$w^2 + x^2 + y^2 + z^2 \leq 1000$$

since the conclusion is obvious otherwise.

Define

$$f(w, x, y, z) = \sum_{\text{cyc}} \ln(w^2 + 1)$$

and

$$g(w, x, y, z) = wx + xy + yz + zw + wy + xz - wxyz - 5$$

where we have

$$U = \{(w, x, y, z) \mid w^2 + x^2 + y^2 + z^2 < 1000\}$$

an open ball, whence

$$\bar{U} = \{(w, x, y, z) \mid w^2 + x^2 + y^2 + z^2 \leq 1000\}$$

is a closed ball, which is compact.

First, we need to check that  $\nabla g \neq \mathbf{0}$ . Compute

$$\frac{\partial g}{\partial w} = x + y + z - xyz.$$

If it is the case that all of these are zero, we will derive a contradiction. This would imply that  $wx + wy + wz = wxyz$ . Cyclically adding we get  $2(5 + wxyz) = 4wxyz$ , whence  $wxyz = 5$ . That means  $w(x + y + z) = 5$  so each of  $w, x, y, z$  are the nonzero roots of the quadratic  $t^2 - (w + x + y + z)t + 5 = 0$ , and so they take on at most two distinct values. Now we consider a few cases. If  $w = x = y = z$  then from  $w(x + y + z) = 5$  we must have them all equal to  $\pm\sqrt{5/3}$ , contradicting  $wxyz = 5$ . If  $x = y = z$  but  $w \neq x$ , then  $w + x = w + x + y + z \implies y + z = 0$ , so  $y = z = 0$  which is impossible. Finally, if  $w = x$  and  $y = z$  but  $w \neq y$  then  $w + y = 2(w + y)$ , whence  $w = -y$ . Now we can derive  $5 = w(x + y + z) = -w^2$ , which is impossible. So,  $\nabla g \neq \mathbf{0}$  everywhere.

Now, the extrema on the boundary of  $\bar{U}$  trivially satisfy the inequality, so we only need consider the interior. Evidently we have a  $\lambda$  with

$$\frac{2w}{w^2 + 1} = \lambda(x + y + z - xyz).$$

If  $\lambda = 0$  then  $w = x = y = z = 0$ , which contradicts the given. Otherwise,  $\lambda \neq 0$  whence  $wxyz \neq 0$ . If we let  $s = w + x + y + z$  and  $d = wx + xy + yz + zw + wy + xz$  then we can rearrange this as

$$\frac{2}{\lambda} = \left( \frac{1}{w} + w \right) (x + y + z - xyz)$$

$$\begin{aligned}
&= \frac{w + x + y + z - xyz}{w} - 1 + w(x + y + z) - wxyz \\
&= \frac{s}{w} - 1 + 5 - (xy + yz + zx) \\
\frac{2}{\lambda} - 4 &= \frac{s - (wxy + wyz + wzx + xyz)}{w} = \frac{s - d}{w}.
\end{aligned}$$

Hence we have two cases. Either  $\lambda = \frac{1}{2}$  and  $s = d$ , or

$$w = x = y = z = \frac{s - d}{2/\lambda - 4}$$

In the latter case, we obtain  $6w^2 = w^4 + 5$ , which implies that the constant values are one of  $\pm 1, \pm\sqrt{5}$ . Check that none of these cases give  $f(w, x, y, z) < 16$ .

Otherwise, we derive that

$$w + x + y + z = wxy + xyz + yzw + zwx \tag{1}$$

At this point we have used all the constraints. That means Lagrange Multipliers is telling us that any point  $(w, x, y, z)$  which also satisfies (1) is possibly a local extrema; in other words, one can check that any  $\mathbf{x}$  satisfying (1) and  $wx + xy + yz + zx + wy + xz = 5 + wxyz$  also satisfies  $\nabla f(\mathbf{x}) = \lambda \nabla g(\mathbf{x})$ .

Unfortunately, you can't go farther than this because it turns out that (1) is precisely the equality condition of the original problem. Boo. The end.  $\square$

## §4 Practice Problems

- (USAMO 1978/1) Given that  $a, b, c, d, e$  are real numbers such that

$$\begin{aligned}
a + b + c + d + e &= 8 \\
a^2 + b^2 + c^2 + d^2 + e^2 &= 16,
\end{aligned}$$

determine the minimum value of  $e$ .

- (Mock Inequalities USAMO) Let  $a, b, c, d$  be nonnegative real numbers such that  $a + b + c + d = 4$ . Prove that  $3(a^2 + b^2 + c^2 + d^2) + 4abcd \geq 16$ .
- (MOP 2012) Given that  $a + b + c + d = 4$  for positive reals  $a, b, c, d$ , prove that

$$\frac{1}{a^2} + \frac{1}{b^2} + \frac{1}{c^2} + \frac{1}{d^2} \geq a^2 + b^2 + c^2 + d^2.$$

- (Canada 1999/5) Prove that for nonnegative real numbers  $a, b, c$  with  $a + b + c = 1$  we have

$$a^2b + b^2c + c^2a \leq \frac{4}{27}.$$