The principle of Lagrange

John McCuan

October 11, 2019

Given an open set $\mathcal{U} \subset \mathbb{R}^n$ on which two real valued functions f and g are defined, we consider here the problem of minimizing the function f subject to the constraint $g(\mathbf{x}) = c$, or more explicitly we consider minimizing f on the subset

$$\mathcal{L}_c = \{\mathbf{x} \in \mathcal{U} : g(\mathbf{x}) = c\}.$$

The consideration of this problem is standard in multivariable calculus though the careful statement and proof of a specific result is usually considered beyond the scope of that subject. The Thomas Calculus text, for example, does not state a result, but rather states a "Method of Lagrange." A necessary condition for the existence of an interior point $\mathbf{x}_0 \in \mathcal{U}$ for which $g(\mathbf{x}_0) = c$ and for which

$$f(\mathbf{x}_0) \le f(\mathbf{x}) \qquad \text{for each } \mathbf{x} \in \mathcal{L}_c$$

is not difficult to state:

Theorem 1 (Lagrange) *If $\mathbf{x}_0 \in \mathcal{U} \cap \mathcal{L}_c$ and $f(\mathbf{x}_0) \le f(\mathbf{x})$ for every $\mathbf{x} \in \mathcal{L}_c$, then either $\nabla g(\mathbf{x}_0) = 0$, or there is some real number λ for which*

$$\nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0).$$

Naturally, differentiabiltiy of f and g is required. A heuristic explanation for why the result holds is also relatively easy to give, at least when $n = 2$. Let us give such an explanation.

The first ingredient is a precise qualitative understanding of the behavior of a differentiable function at a point $\mathbf{x}_0$ in terms of its rate of change in various directions, i.e., directional derivatives. When $n = 2$, there is a **tangent plane** at x_0. One must be careful here. Especially a student of elementary calculus must be careful here. I do

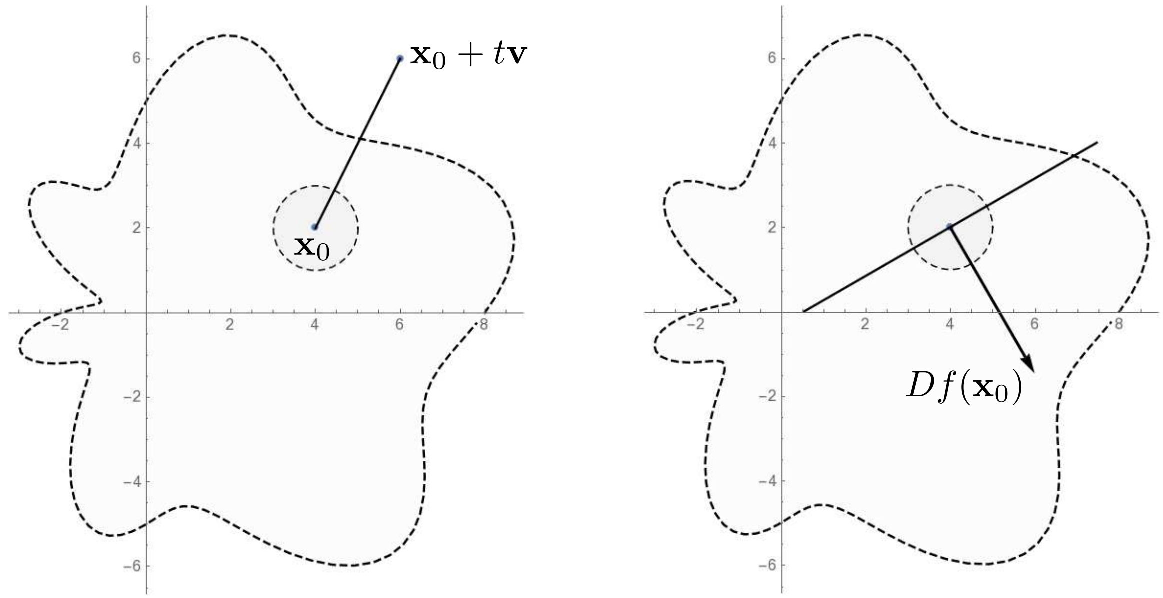

Figure 1: the tangent plane to a domain in $\mathbb{R}^2$

not mean the tangent plane to the graph of the function, which is a plane in $\mathbb{R}^3$, but rather a plane to *the domain* of the function. For various reasons, this notion presents itself as rather abstract and, perhaps, esoteric. Let us try to be rather precise and take some time to understand it.

Tangent plane to a domain

Let $f : \mathcal{U} \to \mathbb{R}^1$ be a real valued function defined on an open set $\mathcal{U} \subset \mathbb{R}^2$ with a point $\mathbf{x}_0 \in \mathcal{U}$. The set $\mathcal{U}$ may not be the entire plane $\mathbb{R}^2$, as indicated on the left in Figure 1, but we can take a line segment starting at $\mathbf{x}_0 \in \mathcal{U}$ in the direction $\mathbf{v}$ and find a point $\mathbf{x}_0 + t\mathbf{v}$ which is outside of $\mathcal{U}$. Such a point is still in the tangent plane to $\mathcal{U}$ at the point $\mathbf{x}_0$. I hope it is clear that, as we look at Figure 1, three distinct sets are under consideration:

1. The *ambient space* $\mathbb{R}^2$,
2. the *domain* $\mathcal{U}$ which is a subset of $\mathbb{R}^2$, and
3. the **tangent plane** at $\mathbf{x}_0$ to $\mathcal{U}$.

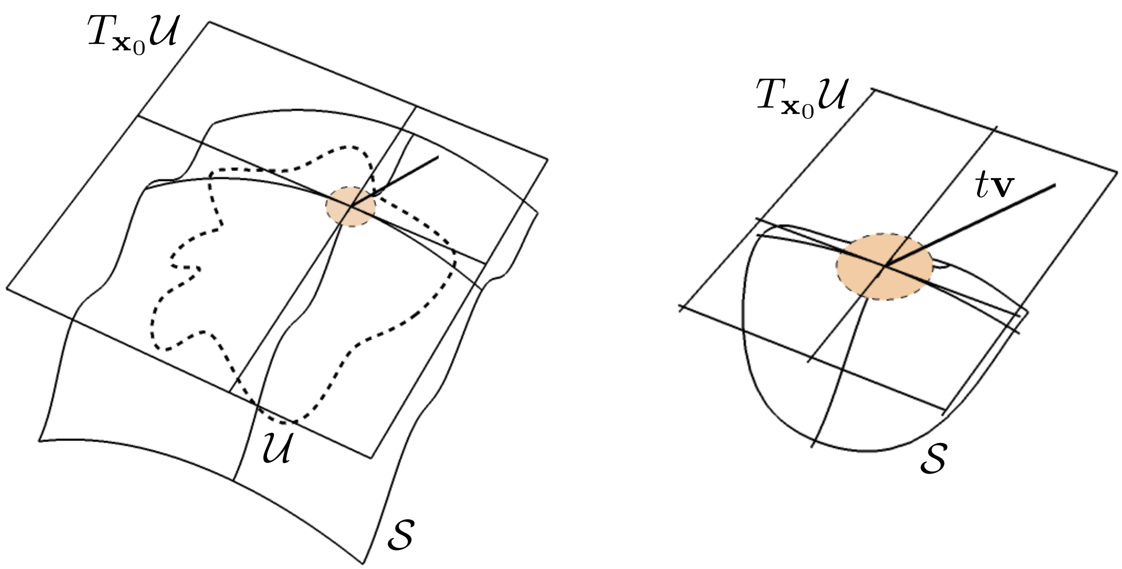

Figure 2: the tangent plane to a domain $\mathcal{U}$ in a surface $\mathcal{S}$

The tangent plane coincides with the ambient space $\mathbb{R}^2$ as a set, but it is natural, and it can be useful, to think of these two conincident sets as distinct.

A second way which may be helpful to distinguish the ambient space $\mathbb{R}^2$ and the tangent plane $T_{\mathbf{x}_0}\mathcal{U}$ to $\mathcal{U}$ at $\mathbf{x}_0$ is by introducing distinct coordinates on each set. One can see the origin in $\mathbb{R}^2$ indicated in Figure 1, and the coordinates of the point $\mathbf{x}_0$ in the figure are given by $\mathbf{x}_0 = (4, 2)$. In the tangent space to $\mathcal{U}$, we may introduce coordinates in which $\mathbf{x}_0$ is the origin, and the coordinates of the displacement $t\mathbf{v} \in T_{\mathbf{x}_0}\mathcal{U}$ are given by $t\mathbf{v} = (2, 4)$. Thus, the ambient space is a space of *points* while the tangent space is a space of *displacements* from points. In particular, $T_{\mathbf{x}_0}\mathcal{U}$ is the space of displacements $\mathbf{v}$ from the point $\mathbf{x}_0$. A point $\mathbf{x}_0 + t\mathbf{v}$ obtained by displacement from $\mathbf{x}_0 \in \mathcal{U}$ is in the ambient space $\mathbb{R}^2$ but may or may not be in $\mathcal{U}$ depending on the displacement $t\mathbf{v}$.

Just in case the distinction we are making between the tangent plane to $\mathcal{U}$ and the ambient space $\mathbb{R}^2$ is not clear, we can further illustrate the distinction by modifiying the domain and its ambient space in a dramatic way. In Figure 2 we have warped the ambient space $\mathbb{R}^2$ of Figure 1 so that it curves downward in all directions at $\mathbf{x}_0$ away from the tangent plane. The open set $\mathcal{U}$ follows the bending of the ambient space, but the tangent plane remains essentially in place. Now the distinction between the ambient space $\mathcal{S}$ and the tangent plane $T_{\mathbf{x}_0}\mathcal{U}$ is clear. On the right we have zoomed in for a close up near the point $\mathbf{x}_0$. The ambient space is now a surface $\mathcal{S}$ in $\mathbb{R}^3$, and the coordinates of $\mathbf{x}_0$ may be taken to be $\mathbf{x}_0 = (4, 2, 0)$. The point $\mathbf{x}_0 + t\mathbf{v}$ is only in the ambient space $\mathcal{S}$ when $t = 0$. We have indicated the unit disk in the tangent space on the right and on the left. It is still natural to use the coordinate expressions $t\mathbf{v} = (2, 4)$ and $\mathbf{v} = (1, 2)/\sqrt{5}$ in the tangent space. But an adjustment of coordinates will now be required to add displacements: The particular point $\mathbf{x}_0 + t\mathbf{v} = (4, 2, 0) + (2, 4, 0) = (6, 6, 0)$ is in a plane in $\mathbb{R}^3$ which coincides with the tangent plane.

Now, let us return to our focus on the tangent plane to $\mathcal{U}$ in Figure 1. The gradient vector gives a direction of maximum increase of the values of f. The normal line to this vector divides the tangent plane into two half spaces. Let us take $\mathbf{u}_1 = Df/|Df|$ and $\mathbf{u}_2$ the rotation by $\pi/2$ counterclockwise of $\mathbf{u}_1$. Then every unit vector in the tangent plane may be expressed as

$$\mathbf{v} = \cos\theta\, \mathbf{u}_1 + \sin\theta\, \mathbf{u}_2$$

where θ is some angle. Then

$$D_{\mathbf{u}_1} f(\mathbf{x}_0) = Df(\mathbf{x}_0) \cdot \mathbf{u}_1 = |Df(\mathbf{x}_0)| > 0 \qquad \text{and} \qquad D_{\mathbf{u}_2} f(\mathbf{x}_0) = Df(\mathbf{x}_0) \cdot \mathbf{u}_2 = 0.$$

More generally, the directional derivative in the direction $\mathbf{v} = \cos\theta\, \mathbf{u}_1 + \sin\theta\, \mathbf{u}_2$ is given by

$$D_{\mathbf{v}} f(\mathbf{x}_0) = \cos\theta Df(\mathbf{x}_0) \cdot \mathbf{u}_1 + \sin\theta Df(\mathbf{x}_0) \cdot \mathbf{u}_2 = \cos\theta |Df(\mathbf{x}_0)|.$$

This gives us a very precise understanding of the directional derivatives $D_{\mathbf{v}} f(\mathbf{x}_0)$. Moving around the unit circle in $T_{\mathbf{x}_0}\mathcal{U}$, there is exactly one maximum direction along $\mathbf{u}_1$; there are exactly two directions yielding zero directional derivative, namely $\pm\mathbf{u}_2$. Thus, the line in $T_{\mathbf{x}_0}\mathcal{U}$ determined by the direction $\mathbf{u}_2$ divides the tangent plane into two half planes. The half plane into which $Df(\mathbf{x}_0)$ points consists of directions in which f increases to first order. The other half plane consists of directions in which the corresponding directional derivative of f is negative. Let us call this line the **dividing line**:

$$Z = \{\mathbf{x}_0 + t\mathbf{u}_2 : t \in \mathbb{R}\} \subset T_{\mathbf{x}_0}\mathcal{U}.$$

There is one situation in which our discussion is not valid. That is when the gradient vanishes. We can consider this case in more detail later, but for now simply note that if $Df(\mathbf{x}_0) = (0, 0)$, the conclusion of Theorem 1 is easily seen to hold with $\lambda = 0$.

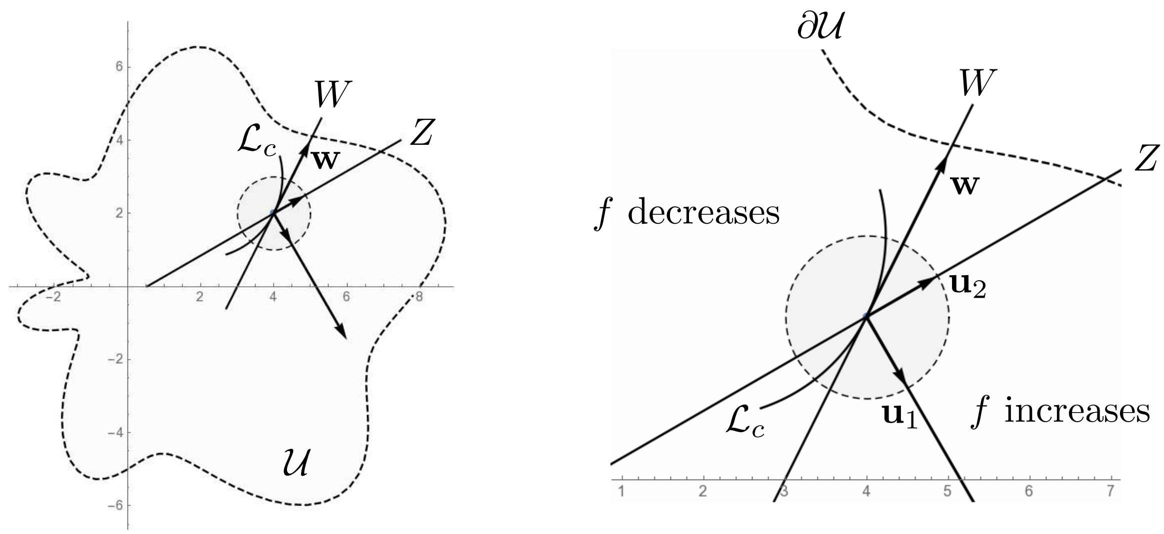

Figure 3: a level curve $\mathcal{L}_c = \{\mathbf{x} : f(\mathbf{x}) = c\}$ when $Dg(\mathbf{x}_0)$ and $Dg(\mathbf{x}_0)$ are not parallel

Now, we are in a position to see, heuristically, why the theorem holds. If $Dg(\mathbf{x}_0) \neq 0$, then we expect the level set $\mathcal{L}_c$ near $\mathbf{x}_0$ will be a differentiable curve as shown in Figure 3. Moreover, if $Dg(\mathbf{x}_0)$ is not parallel to $Df(\mathbf{x}_0)$, then the curve will cross the dividing line Z at $\mathbf{x}_0$ and determine a direction $\mathbf{w}$ which is not tangent to Z. Let us assume we can parameterize the level set locally by a function $\mathbf{r} : (-\epsilon, \epsilon) \to \mathcal{U}$ with $\mathbf{r}(0) = 0$ and $\mathbf{r}'(0) = \mathbf{w}$. (Here ϵ is just some positive number.) Then the tangent vector to $\mathcal{L}_c$ is $\mathbf{w} = \mathbf{r}'(0)$ as indicated in Figure 3, and the tangent line to $\mathcal{L}_c$ at $\mathbf{x}_0$ is

$$W = \{\mathbf{x}_0 + t\mathbf{r}'(0) : t \in \mathbb{R}\} \subset T_{\mathbf{x}_0}\mathcal{U}.$$

Again, we have zoomed in to the point $\mathbf{x}_0$ on the right in Figure 3, and in this case, it's clear that at least one of the directions along $\mathcal{L}_c$ will enter the region where f has values smaller than $f(\mathbf{x}_0)$. To be precise, we can find $\mathbf{r}$ so that $\mathbf{r}'(0) = \mathbf{w} = a\mathbf{u}_1 + b\mathbf{u}_2$ with $a < 0$ and

$$\frac{d}{dt} f(\mathbf{r}(t))\Big|_{t=0} = Df(\mathbf{x}_0) \cdot \mathbf{r}'(0) = Df(\mathbf{x}_0) \cdot \mathbf{w} = D_{\mathbf{w}} f(\mathbf{x}_0) = a|Df(\mathbf{x}_0)| < 0.$$

This means that for $0 < t < \epsilon$ with t small

$$f(\mathbf{r}(t)) < f(\mathbf{x}_0).$$

Since $\mathbf{r}(t) \in \mathcal{L}_c$, this means $\mathbf{x}_0$ cannot be a (local) minimum point.

Making this heuristic argument rigorous

The deficiency of the argument above centers mainly on the assumption that $\mathcal{L}_c$ is a differentiable curve. More broadly the distinction between a differentiable curve as a set and as a mapping $\mathbf{r} : (-\epsilon, \epsilon) \to \mathbb{R}^2$ is usually not addressed carefully in an elementary calculus course. Thus, there is a substantial gap in rigor between the assumption that $\mathbf{x}_0$ is a minimum point in $\mathcal{L}_c$ and the existence of a corresponding curve parameterized by $\mathbf{r} : (-\epsilon, \epsilon) \to \mathcal{U}$ leading to lower values of f taken on $\mathcal{L}_c$ unless $Dg(\mathbf{x}_0)$ and $Df(\mathbf{x}_0)$ are parallel.

These questions are usually addressed in a course on *differential geometry* and require an application of the *inverse function theorem* or the closely related *implicit function theorem*, neither of which is usually covered in elementary calculus. I will not cover all this material here. But I will state, for the benefit of precocious and interested students of elementary calculus, a result that can be used to make our heuristic argument rigorous.

It should first be noted that the level set $\mathcal{L}_c$ need not be a differentiable curve. If, for example, $g(x, y) = (x-4)^2 - (y-2)^2$ and $c = 0$, it is easy to see $\mathcal{L}_0$ consists of two intersecting straight lines. This situation, in contrast to the hopeful representation of Figure 3, does not give that $\mathcal{L}_0$ is a differentiable curve in any neighborhood of $(4, 2)$. In fact, the set $\mathcal{L}_c$ may not even *contain* a differentiable curve. One runs into this problem if $g(x, y) = -(x-4)^2 - (y-2)^2$ and, again, $c = 0$. These observations explain the appearance of the alternative case $Dg(\mathbf{x}_0) = 0$ in the statement of Theorem 1. The standard result invoked to prove the main case is the following:

Theorem 2 *If $g : \mathcal{U} \to \mathbb{R}^1$ satisfies $g \in C^1(\mathcal{U})$ and $Dg(\mathbf{x}_0) \neq 0$ for some $\mathbf{x}_0 \in \mathcal{L}_c \subset \mathcal{U}$, then there is some open ball $B_r(\mathbf{x}_0)$ such that $\mathcal{L}_c \cap B_r(0)$ is a* **regular curve***, that is, for each point $\mathbf{p} \in \mathcal{L}_c \cap B_r(0)$, there is some $\epsilon > 0$ and some C^1 function $\mathbf{r} : (-\epsilon, \epsilon) \to \mathcal{L}_c$ such that $\mathbf{r}(0) = \mathbf{p}$ and $\mathbf{r}'(t) \neq 0$ for $-\epsilon < t < \epsilon$. Moreover, if we take a particular parameterization, then there is some $\rho > 0$ with $0 < \rho < r$ and some a and b such that $-\epsilon < a < 0 < b < \epsilon$ such that $\mathcal{L}_c \cap B_\rho(0) = \{\mathbf{r}(t) : a < t < b\}$.*

As in the discussion above, $\mathcal{U}$ is an open set in $\mathbb{R}^2$. Recall that the open ball is defined by $B_r(\mathbf{x}_0) = \{\mathbf{x} : |\mathbf{x} - \mathbf{x}_0| < r\}$. The condition $g \in C^1(\mathcal{U})$ means the first partial derivatives of g exist and are continuous on $\mathcal{U}$. The C^1 condition on $\mathbf{r}$ means $\mathbf{r}'$ exists and is continuous on $(-\epsilon, \epsilon)$.

Once Theorem 2 is established, the connection between the values of points in $\mathcal{L}_c$ near $\mathbf{x}_0$ and the directional derivative in the tangent direction $\mathbf{w} = \mathbf{r}'(0)$ at $\mathbf{x}_0$ follows easily as described in the previous section.

Higher dimensions $n > 2$

Some new ideas are involved in the discussion of the higher dimensional cases, but the basic ideas still work. Let us consider $n = 3$. As above, we may assume

$$Df(\mathbf{x}_0) \neq 0 \qquad \text{and} \qquad Dg(\mathbf{x}_0) \neq 0.$$

The condition on $Df(\mathbf{x}_0)$ serves to separate $T_{\mathbf{x}_0}\mathcal{U}$ (which is now a three dimensional space) into two half spaces H^+, into which $Df(\mathbf{x}_0)$ points, and H^-. Displacements into H^+ correspond to increasing values of f. Displacements into H^- correspond to decreasing values of f. There is now an entire plane of directions in $T_{\mathbf{x}_0}\mathcal{U}$, separating H^- and H^+, where the directional derivatives vanish:

$$D_{\mathbf{v}}f(\mathbf{x}_0) = 0 \qquad \text{whenever} \qquad \mathbf{v} \cdot Df(\mathbf{x}_0) = 0.$$

We have then a **dividing plane**

$$Z = \{\mathbf{v} \in T_{\mathbf{x}_0}\mathcal{U} : \mathbf{v} \cdot Df(\mathbf{x}_0) = 0\}.$$

Theorem 3 *If $g : \mathcal{U} \to \mathbb{R}^1$ satisfies $g \in C^1(\mathcal{U})$, where $\mathcal{U}$ is an open set in $\mathbb{R}^3$, and $Dg(\mathbf{x}_0) \neq 0$ for some $\mathbf{x}_0 \in \mathcal{L}_c \subset \mathcal{U}$, then there is some open ball $B_r(\mathbf{x}_0)$ such that $\mathcal{L}_c \cap B_r(0)$ is a* **regular surface**. *In particular, if we take $r > 0$ small enough, then there is an open set $\mathcal{V} \subset \mathbb{R}^2$ and a C^1 parametric (surface) mapping $X : \mathcal{V} \to \mathbb{R}^3$ such that $X = X(u, v)$ satisfies*

$$X_u \times X_v \neq 0$$

and

$$\mathcal{L}_c \cap B_r(\mathbf{x}_0) = \{X(u, v) : (u, v) \in \mathcal{V}\}.$$

If $Dg(\mathbf{x}_0)$ is not parallel to $Df(\mathbf{x}_0)$, then there will be a tangent direction $\mathbf{w}$ in $T_{\mathbf{x}_0}\mathcal{U}$ which points into H^-. There will also be a parameterized curve $\mathbf{r} : (-\epsilon, \epsilon) \to \mathcal{L}_c$ (taking values in the surface $\mathcal{L}_c$) such that $\mathbf{r}(0) = \mathbf{x}_0$ and $\mathbf{r}'(0) = \mathbf{w}$. Again, the values $f(\mathbf{r}(t))$ for t small and positive will be smaller than $f(\mathbf{x}_0)$ because $\mathbf{r}(t) \in \mathcal{L}_c$ and

$$\frac{d}{dt} f(\mathbf{r}(t))\Big|_{t=0} = Df(\mathbf{x}_0) \cdot \mathbf{w} < 0.$$

For $n > 3$ the **dividing hyperplane**

$$Z = \{\mathbf{v} \in T_{\mathbf{x}_0}\mathcal{U} : \mathbf{v} \cdot Df(\mathbf{x}_0) = 0\}$$

is an $n - 1$ dimensional subspace of $T_{\mathbf{x}_0}\mathcal{U} \approx \mathbb{R}^n$ separating the increasing directions for f from the decreasing ones. There is a higher dimensional analogue of Theorem 3, and the argument goes pretty much as before.

Review/statement of the proof

Let $\mathcal{U} \subset \mathbb{R}^n$ be an open set, and let f and g be C^1 functions defined on $\mathcal{U}$. Finally, let

$$\mathcal{L}_c = \{\mathbf{x} \in \mathcal{U} : g(\mathbf{x}) = c\}.$$

Theorem 4 (Lagrange) *If $\mathbf{x}_0 \in \mathcal{U} \cap \mathcal{L}_c$ and there is some $r > 0$ such that $f(\mathbf{x}_0) \leq f(\mathbf{x})$ for every $\mathbf{x} \in \mathcal{L}_c \cap B_r(\mathbf{x}_0)$, then either $\nabla g(\mathbf{x}_0) = 0$, or there is some real number λ for which*

$$\nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0).$$

Proof: If $\nabla f(\mathbf{x}_0) = 0$ or $\nabla g(\mathbf{x}_0) = 0$, then the result clearly holds. Thus, we may assume $\nabla f(\mathbf{x}_0) \neq 0$ and $\nabla g(\mathbf{x}_0) \neq 0$. If $\nabla g(\mathbf{x}_0)$ is **not** parallel to $\nabla f(\mathbf{x}_0)$, i.e., there is no $\lambda \in \mathbb{R}$ such that $\nabla f(\mathbf{x}_0) = \lambda \nabla f(\mathbf{x}_0)$, then in some neighborhood of $\mathbf{x}_0$, the level set $\mathcal{L}_c$ is a C^1 hypersurface and there is a nonzero vector $\mathbf{w} \in T_{\mathbf{x}_0}\mathcal{L}_c$ such that

$$\mathbf{w} \in H^- = \{\mathbf{v} \in T_{\mathbf{x}_0}\mathcal{U} : \mathbf{v} \cdot \nabla f(\mathbf{x}_0) < 0\}.$$

Moreover, there is a C^1 curve $\mathbf{r} : (-\epsilon, \epsilon) \to \mathcal{L}_c$ with $\mathbf{r}(0) = \mathbf{x}_0$ and $\mathbf{r}'(0) = \mathbf{w}$. It follows that

$$\frac{d}{dt} f(\mathbf{r}(t))\Big|_{t=0} = D_{\mathbf{w}} f(\mathbf{x}_0) < 0.$$

Hence there are points $\mathbf{r}(t) \in \mathcal{L}_c$ arbitrarily close to $\mathbf{x}_0$ with $f(\mathbf{r}(t)) < f(\mathbf{x}_0)$. This contradiction shows that $\nabla f(\mathbf{x}_0)$ and $\nabla g(\mathbf{x}_0)$ are parallel. $\square$

Exercise 1 *There is one point, aside from the proofs of Theorems 2 and 3, on which some detail could be added. This point should be accessible to students of elementary calculus with some knowledge of linear algebra. It is claimed, for example in the proof above, that the vector $\mathbf{w} \in T_{\mathbf{x}_0}\mathcal{L}_c$ pointing into H^- exists. Let $\mathbf{w}$ be the projection of $-\nabla f(\mathbf{x}_0)$ onto $T_{\mathbf{x}_0}\mathcal{L}_c$, and show that this choice gives the desired vector. Hint: As a vector in $T_{\mathbf{x}_0}\mathcal{U}$*

$$-\nabla f(\mathbf{x}_0) = (-\nabla f(\mathbf{x}_0) \cdot \mathbf{n})\mathbf{n} + \mathbf{w}$$

can be written uniquely as a component along $\mathbf{n} = \nabla g(\mathbf{x}_0)/|\nabla g(\mathbf{x}_0)|$ and a component orthogonal to $\mathbf{n}$.