Applications

Representation of surfaces

Theorem 1. *Let* $f: \mathbb{R}^N \mapsto \mathbb{R}$, $\boldsymbol{x}_0 \in \mathbb{R}^N$. *Then* $f(\boldsymbol{x}) = f(\boldsymbol{x}_0)$ *is a surface in* $\mathbb{R}^N$. *In particular, if* $\operatorname{grad} f(\boldsymbol{x}_0) \neq \boldsymbol{0}$, *then the tangent plane at* $\boldsymbol{x}_0$ *is*

$$(\operatorname{grad} f)(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0) = 0. \tag{1}$$

Proof. Since $\operatorname{grad} f(\boldsymbol{x}_0) \neq \boldsymbol{0}$, there is at least one x_i such that $\left(\frac{\partial f}{\partial x_i}\right)(\boldsymbol{x}_0) \neq 0$. Thus we can apply implicit function theorem and represent x_i as a function of the other $N-1$ variables. □

Example 2. Find the tangent planes for the sphere $x^2 + y^2 + z^2 = R^2$.

Solution. We check

$$\operatorname{grad} f = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{2}$$

whenever $x^2 + y^2 + z^2 = R^2$. We see that the equation is

$$x_0 x + y_0 y + z_0 z = R^2. \tag{3}$$

Theorem 3. *Consider the curve defined through*

$$f(x, y, z) = 0, \qquad g(x, y, z) = 0. \tag{4}$$

then the equation for the tangent line for the curve is

$$(\operatorname{grad} f)(x_0, y_0, z_0) \cdot \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} = 0 \tag{5}$$

$$(\operatorname{grad} g)(x_0, y_0, z_0) \cdot \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} = 0. \tag{6}$$

Proof. Exercise. □

Lagrange multiplier theory

Recall that when finding optimum of $f: E \subseteq \mathbb{R}^N \mapsto \mathbb{R}$, since usually E is a closed set, we have to consider the following two cases separately:

- E^o: In the interior we solve $\operatorname{grad} f = 0$ to obtain candidates;
- ∂E: So far we have to calculate the values of f on ∂E explicitly.

Clearly this is not satisfactory.

We notice that in optimization problems, the boundary ∂E is usually given through conditions like

$$\boldsymbol{\phi}(\boldsymbol{x}) \geqslant \boldsymbol{0} \tag{7}$$

for some "constraint" function $\boldsymbol{\phi}(\boldsymbol{x}) := \begin{pmatrix} \phi_1(\boldsymbol{x}) \\ \vdots \\ \phi_K(\boldsymbol{x}) \end{pmatrix}$.

We will postpone the dealing of this general situation to a later section when we discuss the Karush-Kuhn-Tucker (KKT) conditions. Here we consider the following problem

$$\max_{\phi(\boldsymbol{x})} f(\boldsymbol{x}) \tag{8}$$

which in optimization literature is usually written as

$$\max f(\boldsymbol{x}) \qquad \text{subject to } \phi(\boldsymbol{x}) = 0. \tag{9}$$

Here $\phi: \mathbb{R}^N \mapsto \mathbb{R}$ is a scalar function. $\phi(\boldsymbol{x}) = 0$ is called a "constraint" of the problem.

Theorem 4. (Lagrange multiplier) *Let $\varnothing \neq U \subseteq \mathbb{R}^N$ be open, let $f, \phi \in C^1$, and let $x_0 \in U$ be such that f has a local maximum or minimum, at x_0 under the constraint $\phi(x) = 0$ and such that $\nabla \phi(x_0) \neq 0$. then there is Lagrange multiplier.*

Proof. As grad $\phi \neq 0$, we can apply the Implicit function theorem. Wlog, assume

$$x_N = X(x_1, \ldots, x_{N-1}). \tag{10}$$

Now $(x_{01}, \ldots, x_{0N-1})$ is a local maximizer/minimizer of the following function

$$F(x_1, \ldots, x_{N-1}) := f(x_1, \ldots, x_{N-1}, X(x_1, \ldots, x_{N-1})). \tag{11}$$

Now applying the necessary condition we have

$$\frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial x_N} \frac{\partial X}{\partial x_i} = 0 \tag{12}$$

for all $i = 1, 2, \ldots, N-1$. Since we have

$$\phi(x_1, \ldots, x_{N-1}, X) = 0 \tag{13}$$

we obtain

$$\frac{\partial \phi}{\partial x_i} + \frac{\partial \phi}{\partial x_N} \frac{\partial X}{\partial x_i} = 0. \tag{14}$$

Recall that $\frac{\partial \phi}{\partial x_N} \neq 0$ so we can solve $\frac{\partial X}{\partial x_i}$ and substitute into the f equation to obtain

$$\frac{\partial f}{\partial x_i} = \left[\left(\frac{\partial \phi}{\partial x_N} \right)^{-1} \frac{\partial f}{\partial x_N} \right] \frac{\partial \phi}{\partial x_i}. \tag{15}$$

If we denote

$$\lambda := \left(\frac{\partial \phi}{\partial x_N} \right)^{-1} \frac{\partial f}{\partial x_N} \tag{16}$$

we conclude that

$$\operatorname{grad} f = \lambda \operatorname{grad} \phi. \tag{17}$$

Thus ends the proof. □

Remark 5. Often a "Lagrange function" is defined:

$$L(\lambda, \boldsymbol{x}) := f(\boldsymbol{x}) - \lambda\, \phi(\boldsymbol{x}). \tag{18}$$

The necessary condition is now stated as $\mathrm{grad}_{\lambda, \boldsymbol{x}} L = \mathbf{0}$.

Exercise 1. Prove that $\mathrm{grad}_{\lambda, \boldsymbol{x}} L = \mathbf{0}$ is a necessary condition for $\boldsymbol{x}_0$ to be a local maximizer/minimizer.

Example 6. Find maximum/minimum of

$$f(x, y) = x\, y \tag{19}$$

on $(x-1)^2 + y^2 = 1$.

Solution. We write the Lagrange function

$$L(\lambda, x, y) = x\, y - \lambda\, [(x-1)^2 + y^2 - 1]. \tag{20}$$

Now we have

$$0 = \frac{\partial L}{\partial x} = y - 2\,(x-1)\,\lambda; \tag{21}$$

$$0 = \frac{\partial L}{\partial y} = x - 2\, y\, \lambda; \tag{22}$$

$$0 = \frac{\partial L}{\partial \lambda} = (x-1)^2 + y^2 - 1. \tag{23}$$

From the first two equations we can cancel λ and obtain

$$y^2 = x\,(x-1). \tag{24}$$

Substituting into the 3rd equation, we get

$$(x-1)^2 + x\,(x-1) - 1 = 0 \Longleftrightarrow 2\, x^2 - 3\, x = 0 \tag{25}$$

and then

$$x = 0, \qquad x = \frac{3}{2}. \tag{26}$$

Correspondingly we have

$$y = 0, \qquad \pm\frac{\sqrt{3}}{2}. \tag{27}$$

Thus we have three candidates: $\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 3/2 \\ \sqrt{3}/2 \end{pmatrix}, \begin{pmatrix} 3/2 \\ -\sqrt{3}/2 \end{pmatrix}$.

Now calculate

$$f(0,0) = 0, \tag{28}$$

$$f\left(\frac{3}{2}, \frac{\sqrt{3}}{2}\right) = \frac{3\sqrt{3}}{4}, \tag{29}$$

$$f\left(\frac{3}{2}, -\frac{\sqrt{3}}{2}\right) = -\frac{3\sqrt{3}}{4}. \tag{30}$$

We see that $\begin{pmatrix} 3/2 \\ \sqrt{3}/2 \end{pmatrix}$ is the maximizer, and $\begin{pmatrix} 3/2 \\ -\sqrt{3}/2 \end{pmatrix}$ is the minimizer.

Example 7. Find maximum of $x_1 \cdots x_n$ satisfying $x_1 + \cdots + x_n = 1$, $x_1, \ldots, x_n \geqslant 0$.

Solution. Here we have the difficulty of n inequality constraints: $x_1 \geqslant 0, \ldots, x_n \geqslant 0$. We will discuss general theory of optimization problems with inequality constraints in a future lecture. On the other hand, for this particular problem we claim that simply solving

$$\max x_1 \cdots x_n \quad \text{subject to } x_1 + \cdots + x_n = 1 \tag{31}$$

is enough.

Let $E := \{\boldsymbol{x} |\ x_1 + \cdots + x_n = 1,\ x_1, \ldots, x_n \geqslant 0\}$. We see that this is a bounded closed set and therefore the continuous function $x_1 \cdots x_n$ must reach its maximum in E. It is easy to see that at the maximum, it must be $x_1 > 0, \ldots, x_n > 0$, which means the maximizer at least corresponds to a local maximizer for the problem

$$\max x_1 \cdots x_n \quad \text{subject to } x_1 + \cdots + x_n = 1 \tag{32}$$

Define the Lagrange function

$$L(\lambda, x_1, \ldots, x_n) := x_1 \cdots x_n - \lambda (x_1 + \cdots + x_n - 1). \tag{33}$$

Taking partial derivatives we have

$$\begin{aligned} 0 = \frac{\partial L}{\partial x_1} &= x_2 \cdots x_n - \lambda, && (34) \\ \vdots \quad &\vdots \quad \vdots \\ 0 = \frac{\partial L}{\partial x_n} &= x_1 \cdots x_{n-1} - \lambda, && (35) \\ 0 = \frac{\partial L}{\partial \lambda} &= x_1 + \cdots + x_n - 1. && (36) \end{aligned}$$

From the first n equations we conclude

$$\frac{x_1 \cdots x_n}{x_i} = \lambda \tag{37}$$

for all $i = 1, 2, \ldots, n$ which gives $x_1 = \cdots = x_n$.[1] Now activating the last equation $x_1 + \cdots + x_n - 1 = 0$ we see that the only candidate for maximizer is $\begin{pmatrix} 1/n \\ \vdots \\ 1/n \end{pmatrix}$ with $f(x_1, \ldots, x_n) = 1/n^n$. Since it is the only candidate, it has to be the maximizer, and the maximum is $1/n^n$.

Problem 1. Develop the Lagrange multiplier theory for multiple constraints: $\phi_1(\boldsymbol{x}) = \cdots = \phi_K(x) = 0$.

1. The other possibility is that one of x_i is 0. But then we know it cannot be the maximizer.