Theory sheet 12

Constrained extrema

Let $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ be two functions.

Definition 1. *We say that a point $\boldsymbol{x}_0 \in \mathbb{R}^n$ is a local maximum (or minimum) of a function under constraint $g(\boldsymbol{x}) = 0$, if $f(\boldsymbol{x}_0) \geq f(\boldsymbol{x})$ (or $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x})$) for all $\boldsymbol{x} \approx \boldsymbol{x}_0$ such that $g(\boldsymbol{x}) = 0$.*

Here are a few remarks:

- The opposite of "constrained" is "free": *free extrema* means "extrema without constraints", that is, over the entire $\mathbb{R}^n$.
- Note the difference with the notion of local maximum/minimum: there we need to compare $f(\boldsymbol{x}_0)$ with values of f at all neighbouring points $\boldsymbol{x}$, whereas here $f(\boldsymbol{x}_0)$ is only compared against those points $\boldsymbol{x}$, where $g(\boldsymbol{x}) = 0$ is satisfied.
- In particular, $\boldsymbol{x}_0$ should itself satisfy $g(\boldsymbol{x}_0) = 0$.
- Why zero in "$g(\boldsymbol{x}) = 0$"? Just convention! Any constraint of the form $g(\boldsymbol{x}) = c$ may be rewritten as $\tilde{g}(\boldsymbol{x}) = 0$ with a new function $\tilde{g}(\boldsymbol{x}) = g(\boldsymbol{x}) - c$.
- In other words, we are studying extrema under equality constraints. There is also a similar theory of the extrema under inequality constraints $g(\boldsymbol{x}) \geq 0$, but we shall not treat it in this course.
- Geometric interpretation: since $G = \{\boldsymbol{x} : g(\boldsymbol{x}) = 0\} \subset \mathbb{R}^n$ is some surface, the constrained extrema of f is just the free extrema of f restricted to this surface. That is, free extrema of $f : G \to \mathbb{R}$ instead of $f : \mathbb{R}^n \to \mathbb{R}$.

Natural examples of a constrained extrema problem come from utility theory, where we usually have budget constraints (amount of money is finite).

Another natural example: maximize the area of a rectangle if its perimeter is fixed. Or maximize the volume of a box if its surface area is fixed.

Below, we shall discuss two distinct approaches for solving this problem:

- **substitution method:** treat the constraint $g(\boldsymbol{x}) = 0$ as an equation, using which we can express some variables in terms of the other and plug it into f; schematically,

$$g(x_1, x_2) = 0 \overset{\text{solve}}{\Longrightarrow} x_2 = h(x_1) \Longrightarrow \text{ optimize new function } \tilde{f}(x_1) = f(x_1, h(x_1)).$$

- **Lagrange multipliers method:** construct a new function L of $n + 1$ variables (introducing an artificial variable), called the Lagrangian, such that the constrained extrema of f are the free extrema of L, and optimize L as before.

Reminder on free extrema

Recall that $\boldsymbol{x}_0$ is a candidate extremal point of f if $\operatorname{grad} f(\boldsymbol{x}_0) = \mathbf{0}$. To determine if $\boldsymbol{x}_0$ is indeed an extremal point, we need to check the type of the second order differential $d^2 f(\boldsymbol{x}_0)$ or of the Hessian $H_f(\boldsymbol{x}_0)$.

Example

In this section we use the substitution method to solve the following problem: maximize

$$B(q_1, q_2) = 55q_1 + 70q_2 - 2q_1^2 - 3q_1q_2 - 3q_2^2$$

under constraint

$$q_1 + q_2 - 13 = 0.$$

First, we resolve the constraint for any of the two variables:

$$q_2 = 13 - q_1,$$

then we plug this into B:

$$\begin{aligned}\tilde{B}(q_1) = B(q_1, 13 - q_1) &= 55q_1 + 70(13 - q_1) - 2q_1^2 - 3q_1(13 - q_1) - 3(13 - q_1)^2 \\ &= 403 + 24q_1 - 2q_1^2.\end{aligned}$$

Then, we maximize $\tilde{B}$ as usual:

$$\tilde{B}'(q_1) = 0 \implies 24 - 4q_1 = 0 \implies q_1 = 6.$$

It remains to find q_2:

$$q_2 = 13 - q_1 = 7.$$

Since $\tilde{B}'' = -4 < 0$, this solution is indeed a maximum.

Necessary conditions for extremum under constraints

The following theorem is given without proof:

Theorem 1. *If $\boldsymbol{x}_0$ is an extremum point of f under constraint $g(\boldsymbol{x}) = 0$, then* $\operatorname{grad} f$ *and* $\operatorname{grad} g$ *are codirectional or proportional. That is, there exists a number $\lambda \in \mathbb{R}$ such that*

$$\operatorname{grad} f(\boldsymbol{x}_0) = \lambda \operatorname{grad} g(\boldsymbol{x}_0).$$

- $\operatorname{grad} f(\boldsymbol{x}_0) = \lambda \operatorname{grad} g(\boldsymbol{x}_0)$ means that both vectors point in the same direction if $\lambda > 0$ and that they point in the opposite directions if $\lambda < 0$.
- The previous theorem is only a necessary condition, not sufficient.
- Second order condition is not applicable here.

The above theorem tells us that in order to optimize f under constraint given by g, we need to solve $n+1$ equations

$$f'_{x_1} = \lambda g'_{x_1}, \quad \dots, \quad f'_{x_n} = \lambda g'_{x_n}, \quad g(x_1, \dots, x_n) = 0$$

on $n+1$ variables

$$x_1, \quad \dots, \quad x_n, \quad \lambda.$$

Lagrangian

The idea of Lagrange multipliers method is to rewrite the condition

$$\operatorname{grad} f(\boldsymbol{x}) = \lambda \operatorname{grad} g(\boldsymbol{x})$$

as follows:

$$\boxed{\operatorname{grad} L(\boldsymbol{x}, \lambda) = 0, \quad \text{where} \quad L(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) - \lambda\, g(\boldsymbol{x}).}$$

Definition 2. *The Lagrangian of the optimization problem for f with constraint given by g is a function of $n+1$ real variables*

$$L(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) - \lambda g(\boldsymbol{x}).$$

Remark 1. *When we write $\operatorname{grad} L(\boldsymbol{x}, \lambda)$, do we mean gradient with respect to $\boldsymbol{x}$, or with respect to $\boldsymbol{x}$ and λ? The answer: it's not important, because the two gradients differ by one (last) coordinate, the derivative with respect to λ, which gives*

$$\frac{\partial L}{\partial \lambda} = \frac{\partial}{\partial \lambda}(f(\boldsymbol{x}) - \lambda g(\boldsymbol{x})) = -g(\boldsymbol{x}).$$

When we set this derivative equal to zero, we simply recover the constraint:

$$g(\boldsymbol{x}) = 0.$$

So, if by grad we mean the gradient with respect to $(\boldsymbol{x}, \lambda)$, then the equation $\operatorname{grad} L = \mathbf{0}$ also encapsulates the constraint $g(\boldsymbol{x}) = 0$. Otherwise, we need state the constraint separately.

Remark 2. *The minus sign in front of $g(\boldsymbol{x})$ in the definition of L is conventional. Since λ is an arbitrary real number, we the problem with $+\lambda$ instead of $-\lambda$ is exactly the same. You may write it as you like.*

Some difference appears in the optimization problems with inequality constraints, which we do not treat in this course anyway.

Example (continued)

Let us now solve the problem

$$\text{maximize} \quad B(q_1, q_2) = 51q_1 + 70q_2 - 2q_1^2 - 3q_1q_2 - 3q_2^2$$

$$\text{under constraint} \quad q_1 + q_2 - 13 = 0$$

using Lagrange multipliers method. First, we build the Lagrangian of this problem:

$$L(q_1, q_2, \lambda) = 51q_1 + 70q_2 - 2q_1^2 - 3q_1q_2 - 3q_2^2 - \lambda(q_1 + q_2 - 13) = 0.$$

Taking the gradient, we obtain the following system of equations:

$$\begin{cases} 51 - 4q_1 - 3q_2 - \lambda = 0, \\ 70 - 3q_1 - 6q_2 - \lambda = 0, \\ -(q_1 + q_2 - 13) = 0. \end{cases}$$

Solving this system as usual, we obtain

$$q_1 = 6, \quad q_2 = 7, \quad \lambda = 10.$$

Remark 3. *Since we introduced λ as an artificial parameter, we don't need to know its value. Hence, we can get rid of it as soon as possible while solving the system.*

Another example

Consider the following optimization problem:

$$\text{maximize} \quad f(x_1, x_2) = x_1x_2 \quad \text{under constraint} \quad 2x_1 + x_2 = 1.$$

Note that the only *free* candidate extrema point $x_1 = x_2 = 0$ of f is a saddle point, but under constraint $2x_1 + x_2 = 1$ the function has a maximum! Let us see this first using the substitution method:

$$2x_1 + x_2 = 1 \implies x_2 = 1 - 2x_1,$$

so we need to find the free extrema of

$$\tilde{f}(x_1) = x_1(1 - 2x_1).$$

Hence,

$$\tilde{f}' = 0 \implies 1 - 4x_1 = 0 \implies x_1 = \frac{1}{4} \implies x_2 = \frac{1}{2}.$$

Since $\tilde{f}'' = -4 < 0$, this point is indeed a maximum.

Economic applications

Let q_1 and q_2 be quantities of some materials, using which we may produce quantity $Q = Q(q_1, q_2)$ of some new object. If $p_1 = 4$ is the price for a unit of material 1 and $p_2 = 3$ is the price for a unit of material 2, and we need $Q = 9$, then the cost minimization problem consists in minimizing the total cost

$$4q_1 + 3q_2$$

under constraint

$$Q(q_1, q_2) = 9.$$

One typical assumption is that Q is of the Cobb-Douglas form. For example,

$$Q(q_1, q_2) = 6q_1^{1/2} q_2^{3/2}.$$

Then the problem is well-posed (that is, we have enough information to begin solving it):

$$\text{minimize} \quad 4q_1 + 3q_2 \quad \text{under} \quad 6q_1^{1/2} q_2^{3/2} = 9.$$

The Lagrangian of this problem is given by

$$L(q_1, q_2, \lambda) = 4q_1 + 3q_2 - \lambda(6q_1^{1/2} q_2^{3/2} - 9).$$

The first order conditions read:

$$\operatorname{grad} L = \mathbf{0} \quad \text{or} \quad \begin{cases} 4 - 3\lambda q_1^{-1/2} q_2^{3/2} = 0 \\ 3 - 9\lambda q_1^{1/2} q_2^{1/2} = 0 \\ -6q_1^{1/2} q_2^{3/2} + 9 = 0. \end{cases}$$

First, we get rid of λ:

$$\lambda = \frac{4}{3} q_1^{1/2} q_2^{-3/2} = \frac{1}{3} q_1^{-1/2} q_2^{-1/2},$$

hence

$$4q_1 = q_2.$$

Plugging this instead of q_2 into the constraint, we obtain

$$-6q_1^{1/2}(4q_1)^{3/2} + 9 = 0 \implies -48q_1^2 + 9 = 0 \implies q_1 = \pm\frac{\sqrt{3}}{4}.$$

Note that the negative solution does not make sense in our interpretation of the problem, so we are left with the following solution:

$$q_1 = \frac{\sqrt{3}}{4} \quad \text{and} \quad q_2 = \sqrt{3}.$$

To find Q, we plug q_1 and q_2 into $Q(q_1, q_2)$:

$$Q = 9.$$

Alternative solution

Let us also solve the same problem by substitution:

$$6q_1^{1/2} q_2^{3/2} = 9 \implies q_1 = \frac{9}{4} q_2^{-3},$$

so, instead of $4q_1 + 3q_2$ under constraint $6q_1^{1/2} q_2^{3/2} = 9$, we need to minimize

$$\tilde{f}(q_1) = 4 \cdot \frac{9}{4} q_2^{-3} + 3q_2.$$

Differentiating, we obtain

$$-27q_2^{-4} + 3 = 0 \implies q_2 = \sqrt{3} \implies q_1 = \frac{\sqrt{3}}{4}.$$

Optimization under many equality constraints

The Lagrange multipliers method also works for optimization problems under many equality constraints:

$$\text{maximize/minimize} \quad f(\boldsymbol{x}) \quad \text{under constraints} \quad g_1(\boldsymbol{x}) = g_2(\boldsymbol{x}) = \cdots = g_m(\boldsymbol{x}) = 0.$$

We just need to build the Lagrangian function as follows:

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{x})$$

where $\boldsymbol{g}(x) = (g_1(\boldsymbol{x}), \ldots, g_m(\boldsymbol{x}))^\top$ or, in other words,

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \sum_{i=1}^{m} \lambda_i g_i(\boldsymbol{x}).$$

Note that L in this case is a function of $n + m$ variables.

The rest of the solution goes as before.

Second order conditions

Second order conditions in constrained problems are more complicated than just checking that Hessian is positive or negative definite. We do not treat these conditions in this course! In particular, this section is not included into the exam.

For those who are interested, we need to check that the Hessian of f *restricted to the constraint tangent space* $F = \{\boldsymbol{v} : \boldsymbol{v}^\top \operatorname{grad} g(\boldsymbol{x}_0) = 0\}$ is positive or negative definite. In other words, that $\boldsymbol{v}^\top H(\boldsymbol{x}_0)\, \boldsymbol{v} < 0$ or > 0 for all non-zero $\boldsymbol{v} \in F$.