Review of Maximization Problem

Xiaoxiao Hu

M.J. Unconstrained Maximization (p.954)

Consider $f : \mathbb{R}^N \to \mathbb{R}$

Definition M.J.1. The vector $\overline{x} \in \mathbb{R}^N$ is a local maximizer of $f(\cdot)$ if there is an open neighborhood of $\overline{x}, A \subset \mathbb{R}^N$, s.t. $f(\overline{x}) \geq f(x)$ for every $x \in A$. If $f(\overline{x}) \geq f(x)$ for every $x \in \mathbb{R}^N$, then $\overline{x}$ is a global maximizer of $f(\cdot)$.

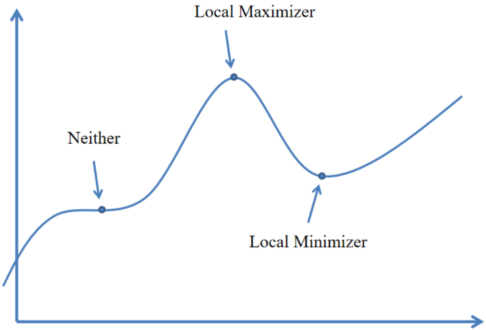
Unconstrained Maximization

Theorem M.J.1. *Suppose that $f(\cdot)$ is differentiable and that $x \in \mathbb{R}^N$ is a local maximizer or local minimizer of $f(\cdot)$. Then $\frac{\partial f(\overline{x})}{\partial x_n} = 0$ for every n, or more concisely*

$$\nabla f(\overline{x}) = \begin{bmatrix} \frac{\partial f(\overline{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\overline{x})}{\partial x_N} \end{bmatrix} = 0.$$

Unconstrained Maximization

Remark. $\nabla f(\overline{x}) = 0$ is only a necessary condition for local maximizer or local minimizer.



Unconstrained Maximization

Theorem M.J.2. *Suppose that the function* $f : \mathbb{R}^N \to \mathbb{R}$ *is twice continuously differentiable* (C^2) *and that* $\nabla f(\overline{x}) = 0.$

(i) *If* $\overline{x} \in \mathbb{R}^N$ *is a local maximizer, then the (symmetric)* $N \times N$ *matrix* $D^2 f(\overline{x})$ *is negative semidefinite.*

(ii) *If* $D^2 f(\overline{x})$ *is negative definite, then* $\overline{x}$ *is a local maximizer.*

Remark. Replacing "negative" by "positive", the same is true for local minimizer.

Unconstrained Maximization

Remark. We rely on the assumption of $z \cdot D^2 f(\overline{x}) z < 0$.

$z \cdot D^2 f(\overline{x}) z \leq 0$ is not enough to guarantee local maximization.

To see this, consider the example, $f(x) = x^3$.

$D^2 f(0)$ is negative semidefinite because $d^2 f(0)/dx^2 = 0$, but $\overline{x} = 0$ is neither a local maximizer nor a local minimizer.

Unconstrained Maximization

Theorem M.J.3. *Any critical point* $\overline{x}$ *(i.e., any* $\overline{x}$ *satisfying* $\nabla f(\overline{x}) = 0$*) of a concave function* $f(\cdot)$ *is a global maximizer of* $f(\cdot)$.

M.K. Constrained Maximization

Case I: Equality Constraints

We first study the maximization problem with M equality constraints, given by (C.M.P.1) below.

$$\max_{x \in \mathbb{R}^N} f(x) \tag{C.M.P.1}$$

$$\text{s.t. } g_1(x) = \overline{b}_1$$

$$\vdots$$

$$g_M(x) = \overline{b}_M$$

Equality Constraints

Constraint Set is

$$C = \{x \in \mathbb{R}^N : g_m(x) = \overline{b}_m \text{ for } m = 1, ..., M\}.$$

Assumption. $N \geq M$ *(Generically, solution doesn't exist if* $M > N$*.)*

Equality Constraints

Theorem M.K.1. *Suppose that the objective and constraint functions of problem (C.M.P.1) are differentiable and that* $\overline{x} \in C$ *is a local constrained maximizer. Assume also that the* $M \times N$ *matrix*

$$\begin{bmatrix} \nabla g_1(\overline{x})^T \\ \vdots \\ \nabla g_M(\overline{x})^T \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1(\overline{x})}{\partial x_1} & \cdots & \frac{\partial g_1(\overline{x})}{\partial x_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_M(\overline{x})}{\partial x_1} & \cdots & \frac{\partial g_M(\overline{x})}{\partial x_N} \end{bmatrix}$$

has rank M. *(This is called **constraint qualification**: It says that the constraints are independent at* $\overline{x}$.*)*

Equality Constraints

Theorem M.K.1 (continued).

Then, there are numbers $\lambda_m \in \mathbb{R}$ *(Not* $\mathbb{R}^+$*), one for each constraint, such that*

$$\frac{\partial f(\overline{x})}{\partial x_n} = \sum_{m=1}^{M} \lambda_m \frac{\partial g_m(\overline{x})}{\partial x_n} \text{ for every } n = 1, ..., N, \qquad \text{(M.K.2)}$$

Or, equivalently,

$$\nabla f(\overline{x}) = \sum_{m=1}^{M} \lambda_m \nabla g_m(\overline{x}). \qquad \text{(M.K.3)}$$

The numbers λ_m are referred to as **Lagrange multipliers**.

How to understand Theorem M.K.1?

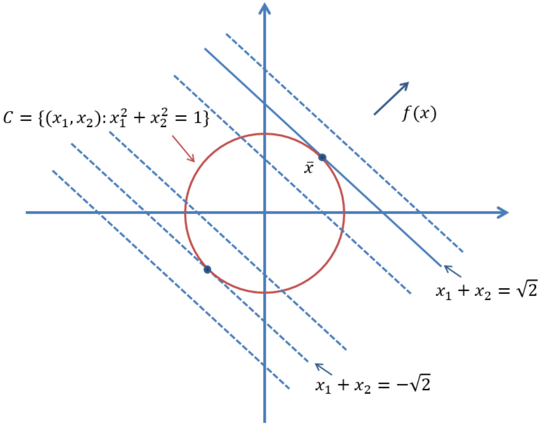
Two-variable, one-constraint Cases

Example M.K.1. Consider the following two-variable, one-constraint example.

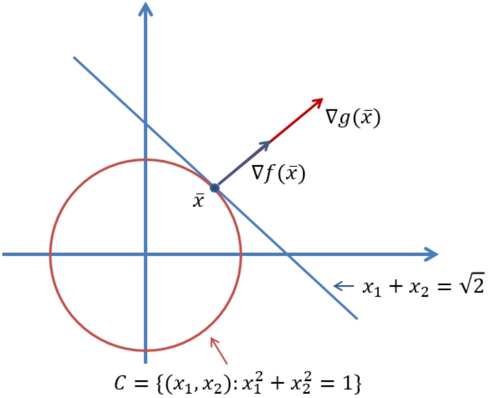
$$\max_{(x_1, x_2) \in \mathbb{R}^2} x_1 + x_2$$

$$\text{s.t. } x_1^2 + x_2^2 = 1$$

Two-variable, One-constraint Example

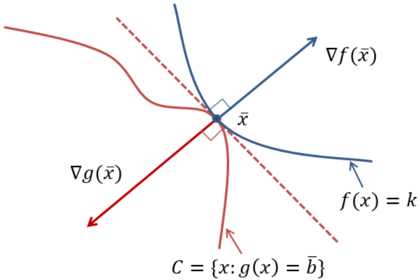


Two-variable, One-constraint Example

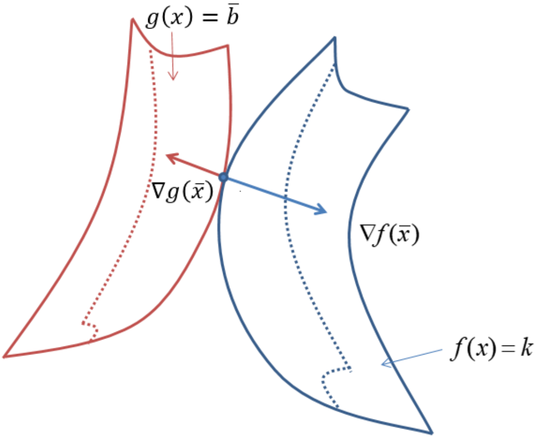


Two-variable, One-constraint Example

More generally, for two-variable, one-constraint cases, the maximum must be obtained where the level set of the objective function is tangent to the constraint set.

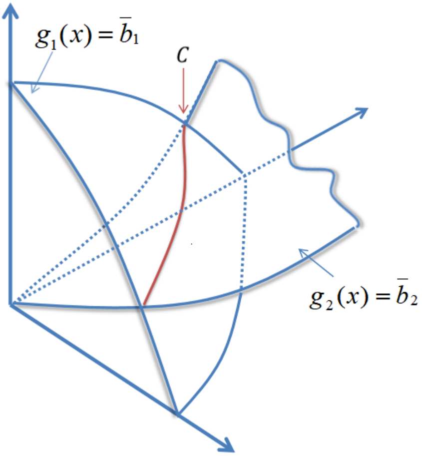


Three-variable, one-constraint cases



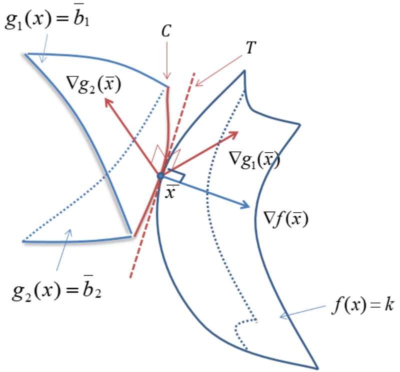
Three-variable, two-constraint cases

Constraint set: $C = \{x \in \mathbb{R}^3 : g_1(x) = \bar{b}_1$ and $g_2(x) = \bar{b}_2\}$



Three-variable, two-constraint cases

Similar to previous cases, the maximum must occur when the level set of the objective function (which is a surface in this case) is tangent to the constraint set.



Three-variable, two-constraint cases

$\nabla f(\overline{x})$, $\nabla g_1(\overline{x})$ and $\nabla g_2(\overline{x})$ are all orthogonal to Line T, implying that they lie on the same plane.

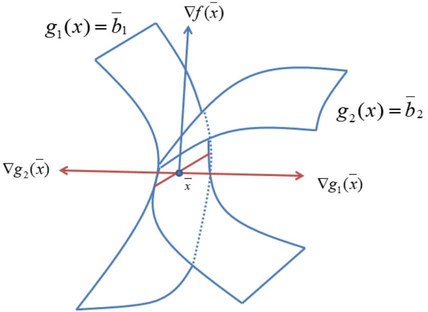
Therefore, $\exists \lambda_1, \lambda_2$ such that $\nabla f(\overline{x}) = \lambda_1 \nabla g_1(\overline{x}) + \lambda_2 \nabla g_2(\overline{x})$, **if $\nabla g_1(\overline{x})$ and $g_2(\overline{x})$ are linearly independent ("Constraint Qualification")**.

More variables

Theorem M.K.1 says that $\nabla f(x)$ lies on the hyperplane spanned by $\nabla g_m(x)$ for $m = 1, ..., M$, **if the constraints** $g_m(x)$ **are linearly independent ("Constraint Qualification")**. The same intuition from previous simple cases apply.

What happens when *Constraint Qualification* fails?

Example. Consider the case $\nabla g_1(\overline{x}) = -\alpha \nabla g_2(\overline{x})$.



Although $\overline{x}$ is a local maximizer, $\nabla f(\overline{x})$ cannot be written as $\lambda_1 \nabla g_1(\overline{x}) + \lambda_2 \nabla g_2(\overline{x})$.

How to use Theorem M.K.1?

Alternative presentation of Theorem M.K.1:

Define *Lagrangian function*:

$$\mathcal{L}(x, \lambda) = f(x) - \sum_{m=1}^{M} \lambda_m (g_m(x) - \overline{b}_m)$$

The constrained maximization problem can be rewritten as the following *unconstrained maximization problem:*

$$\max_{x \in \mathbb{R}^N, \lambda \in \mathbb{R}^M} \mathcal{L}(x, \lambda).$$

Lagrangian Function

First Order Condition (F.O.C.) gives:

$$\frac{\partial f(x)}{\partial x_n} - \sum_{m=1}^{M} \lambda_m \frac{\partial g_m(x)}{\partial x_n} = 0, \text{ for } n = 1, ..., N;$$

$$g_m(x) - \overline{b}_m = 0, \text{ for } m = 1, ..., M.$$

Remark

- In practice, failure of *Constraint Qualification* is rarely a problem. However, you should be alerted and check *Constraint Qualification* if you find the above standard methods problematic.

- If you find no solution, it may be that the maximization problem itself has no solution, or *Constraint Qualification* may fail so that F.O.C is not applicable.

Lagrangian Function

Example. Apply Theorem M.K.1 to solve Example M.K.1:

$$\max_{(x_1,x_2)\in\mathbb{R}^2} x_1 + x_2$$

$$\text{s.t. } x_1^2 + x_2^2 = 1$$

Second Order Condition

- F.O.C. is only a necessary condition for local maximum.

- We also need to check Second Order Condition.

Second Order Condition

If $\overline{x}$ is a local maximizer, then

$$D_x^2 L(\overline{x}, \lambda) = D^2 f(\overline{x}) - \sum_{m=1}^{M} \lambda_m D^2 g_m(\overline{x})$$

is negative semidefinite on the subspace

$$\{z \in \mathbb{R}^N : \nabla g_m(\overline{x}) \cdot z = 0 \text{ for all } m\}.$$

The other direction also applies, i.e., negative definiteness on the subspace implies local maximization.

Second Order Condition

Example. Apply Second Order Condition to the solutions of Example M.K.1.:

$$\max_{(x_1,x_2)\in\mathbb{R}^2} x_1 + x_2$$

$$\text{s.t. } x_1^2 + x_2^2 = 1$$

Method.

- Use the condition in the previous slide directly.
- Use *Bordered Hessian Matrix*

What does λ_m measure?

Claim. λ_m measures the sensitivity of $f(x^*)$ to a small increase in $\overline{b}_m$, i.e., $\lambda_m = \frac{\partial f(x^*(\overline{b}))}{\partial \overline{b}_m}$.

- In class, we consider maximization problem with one constraint only.
- For more constraints, the calculation is similar.

  (See Lecture Notes)

Case II: Inequality Constraints

$$\max_{x \in \mathbb{R}^N} f(x) \tag{C.M.P.2}$$

$$\text{s.t. } g_1(x) \leq \bar{b}_1$$

$$\vdots$$

$$g_M(x) \leq \bar{b}_M$$

Remark. Problem (C.M.P.2) is a simplified version of Problem (M.K.4) in MWG. Here, the coexistence of equality constraints is ignored.

Inequality Constraints

Constraint Set is

$$C = \{x \in \mathbb{R}^N : g_m(x) \leq \bar{b}_m \text{ for } m = 1, ..., M\}.$$

Similar to Theorem M.K.1, we require

Constraint Qualification:

$\nabla g_m(\overline{x})$ with the binding constraints are linearly independent.

Kuhn-Tucker Conditions

Theorem M.K.2 (Kuhn-Tucker Conditions)**.** *Suppose that* $\overline{x} \in C$ *is a local maximizer of problem (C.M.P.2). Assume also that the constraint qualification is satisfied. Then, there are multipliers* $\lambda_m \in \mathbb{R}_+$ *(Not* $\mathbb{R}$*), one for each inequality constraint, such that*

(i) For every $n = 1, ..., N$,

$$\frac{\partial f(\overline{x})}{\partial x_n} = \sum_{m=1}^{M} \lambda_m \frac{\partial g_m(\overline{x})}{\partial x_n} \text{ or } \nabla f(\overline{x}) = \sum_{m=1}^{M} \lambda_m \nabla g_m(\overline{x})$$

Kuhn-Tucker Conditions

Theorem M.K.2 (continued).

(ii) For every $m = 1, ..., M,$

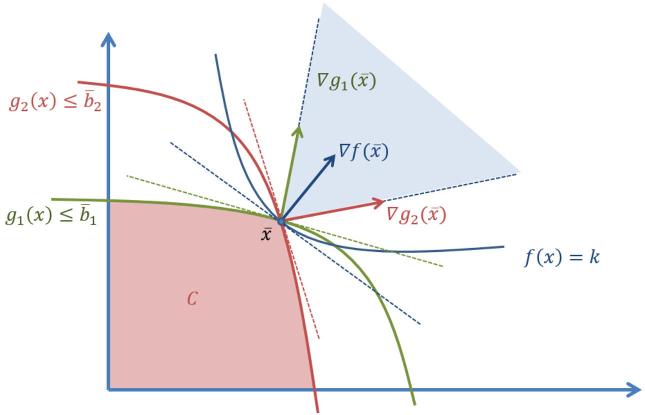
$$\lambda_m(g_m(\overline{x}) - \overline{b}_m) = 0$$

i.e., $\lambda_m = 0$ *for any constraint* k *that doesn't hold with equality.*

Condition (ii) is called "*complementary slackness*" condition: one of the two inequalities $\lambda_m \geq 0$ and $g_m(\overline{x}) \leq \overline{b}_m$ is binding.

Kuhn-Tucker Conditions

Explanation of Kuhn-Tucker Condition (i)



Kuhn-Tucker Conditions

Explanation of Kuhn-Tucker Condition (ii)

When $g_m(\overline{x}) < \overline{b}_m$, the constraint is not binding.

So it doesn't affect the F.O.C locally $\implies \lambda_m = 0$.

How to use Theorem M.K.2?

Define *Lagrangian function*:

$$\mathcal{L}(x,\lambda) = f(x) - \sum_{m=1}^{M} \lambda_m (g_m(x) - \overline{b}_m)$$

How to use Theorem M.K.2?

Kuhn-Tucker conditions give:

$$\frac{\partial f(x)}{\partial x_n} - \sum_{m=1}^{M} \lambda_m \frac{\partial g_m(x)}{\partial x_n} = 0, \text{ for } n = 1, ..., N \quad \text{(FOC for } x_n\text{)}$$

$$g_m(x) - \overline{b}_m \leq 0, \text{ for } m = 1, ..., M \qquad \text{(Constraints)}$$

$$\lambda_m \geq 0, \text{ for } m = 1, ..., M \qquad \text{(Non-negativity of } \lambda\text{)}$$

$$\lambda_m(g_m(x) - \overline{b}_m) = 0, \text{ for } m = 1, ..., M$$

(Complementary slackness)

Second Order Condition

Second order conditions for inequality problems (C.M.P.2) is exactly the same as those for equality problems (C.M.P.1). The only adjustment is that the constraints that count are those that bind, that is, those that hold with equality at the point $\overline{x}$ under consideration.

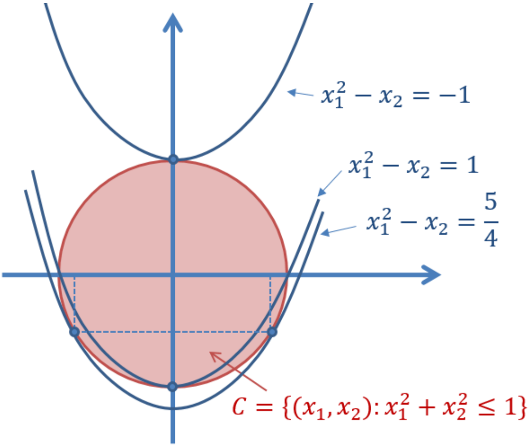
Example of Inequality Constraints

Example M.K.2. Use Theorem M.K.2 to solve the following problem:

$$\max_{(x_1,x_2)\in\mathbb{R}^2} x_1^2 - x_2$$

$$\text{s.t. } x_1^2 + x_2^2 \leq 1$$

Example M.K.2



Adding Non-negativity Constraints

$$\max_{x \in \mathbb{R}^N} f(x)$$

$$\text{s.t. } g_1(x) \leq \overline{b}_1$$

$$\vdots$$

$$g_M(x) \leq \overline{b}_M$$

$$x_1 \geq 0$$

$$\vdots$$

$$x_N \geq 0$$

Adding Non-negativity Constraints

We only need to modify Part (i) of Theorem M.K.2 to

$$\frac{\partial f(\overline{x})}{\partial x_n} \leq \sum_{m=1}^{M} \lambda_m \frac{\partial g_m(\overline{x})}{\partial x_n}, \text{with equality if } \overline{x}_n > 0.$$