

# Review of Maximization Problem

Xiaoxiao Hu

October 21, 2019

## M.J. Unconstrained Maximization (p.954)

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

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

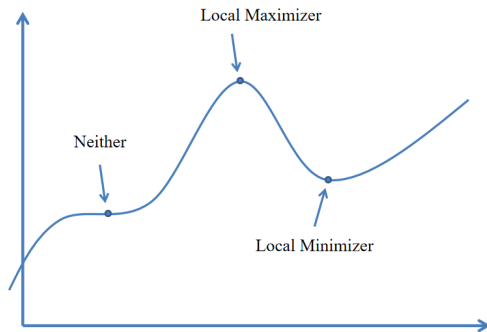
## Unconstrained Maximization

**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  $\frac{\partial f(\bar{x})}{\partial x_n} = 0$  for every  $n$ , or more concisely*

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

## Unconstrained Maximization

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



## Unconstrained Maximization

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

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

*(ii) If  $D^2 f(\bar{x})$  is negative definite, then  $\bar{x}$  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  $z \cdot D^2 f(\bar{x})z < 0$ .

$z \cdot D^2 f(\bar{x})z \leq 0$  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

$\bar{x} = 0$  is neither a local maximizer nor a local minimizer.

## Unconstrained Maximization

**Theorem M.J.3.** *Any critical point  $\bar{x}$  (i.e., any  $\bar{x}$  satisfying  $\nabla f(\bar{x}) = 0$ ) 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.

$$\begin{aligned} & \max_{x \in \mathbb{R}^N} f(x) && \text{(C.M.P.1)} \\ & \text{s.t. } g_1(x) = \bar{b}_1 \\ & \quad \vdots \\ & \quad g_M(x) = \bar{b}_M \end{aligned}$$



## Equality Constraints

**Constraint Set** is

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

**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  $\bar{x} \in C$  is a local constrained maximizer. Assume also that the  $M \times N$  matrix*

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

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

## 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*

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

*Or, equivalently,*

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

The numbers  $\lambda_m$  are referred to as **Lagrange multipliers**.

## How to understand Theorem M.K.1?

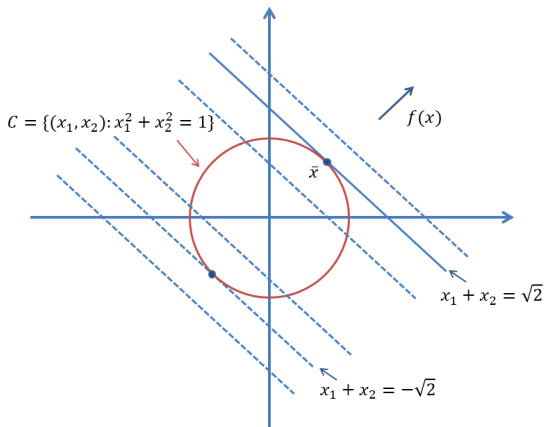
### Two-variable, one-constraint Cases

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

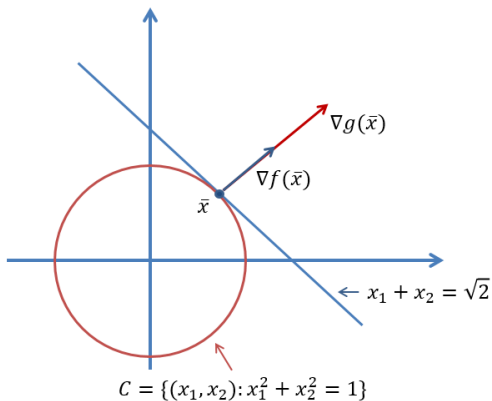
$$\max_{(x_1, x_2) \in \mathbb{R}^2} x_1 + x_2$$

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

## Two-variable, One-constraint Example

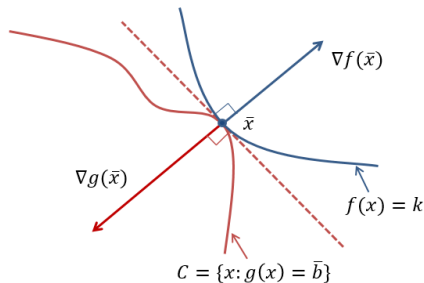


## Two-variable, One-constraint Example

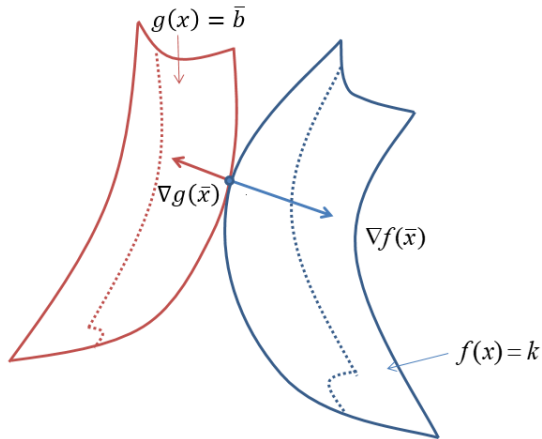


## Two-variable, One-constraint Example

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.



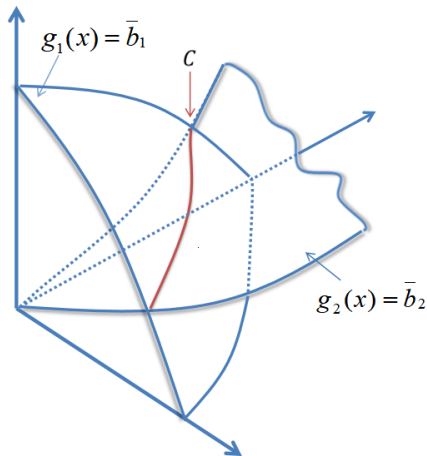
## Three-variable, one-constraint cases





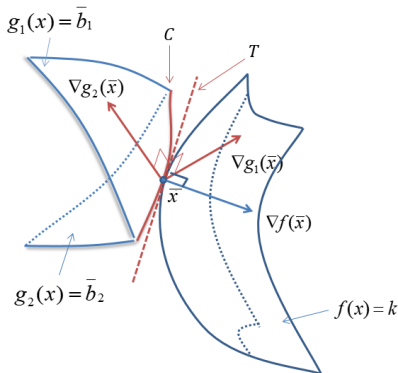
## Three-variable, two-constraint cases

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



## 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

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

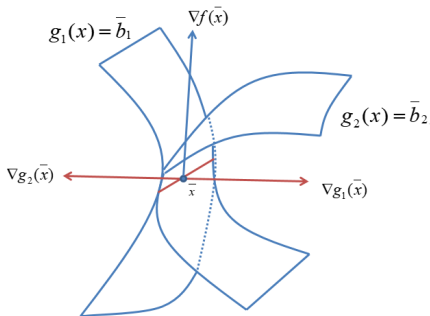
Therefore,  $\exists \lambda_1, \lambda_2$  such that  $\nabla f(\bar{x}) = \lambda_1 \nabla g_1(\bar{x}) + \lambda_2 \nabla g_2(\bar{x})$ ,  
**if  $\nabla g_1(\bar{x})$  and  $\nabla g_2(\bar{x})$  are linearly independent (“Constraint Qualification”).**

## More variables

Theorem [M.K.1](#) says that  $\nabla f(x)$  lies on the hyperplane spanned by  $\nabla g_m(x)$  for  $m = 1, \dots, M$ , **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  $\nabla g_1(\bar{x}) = -\alpha \nabla g_2(\bar{x})$ .



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

## How to use Theorem M.K.1?

Alternative presentation of Theorem M.K.1:

Define *Lagrangian function*:

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

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:

$$\nabla f(x) - \sum_{m=1}^M \lambda_m \nabla g_m(x) = 0$$

$$g_m(x) - \bar{b}_m = 0, \quad m = 1, \dots, M.$$

## 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

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

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

is negative semidefinite on the subspace

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

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](#)..

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

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

**What does  $\lambda_m$  measure?**

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

## Case II: Inequality Constraints

$$\begin{aligned} & \max_{x \in \mathbb{R}^N} f(x) && \text{(C.M.P.2)} \\ & \text{s.t. } g_1(x) \leq \bar{b}_1 \\ & \quad \vdots \\ & \quad g_M(x) \leq \bar{b}_M \end{aligned}$$

*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

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

Similar to Theorem [M.K.1](#), we require

**Constraint Qualification:**

$$\{\nabla g_m : g_m(\bar{x}) = \bar{b}_m, m = 1, \dots, M\}$$

are linearly independent.

## Kuhn-Tucker Conditions

**Theorem M.K.2** (Kuhn-Tucker Conditions). *Suppose that  $\bar{x} \in C$  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  $n = 1, \dots, N$ ,

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

## Kuhn-Tucker Conditions

### Theorem M.K.2 (continued).

(ii) For every  $m = 1, \dots, M$ ,

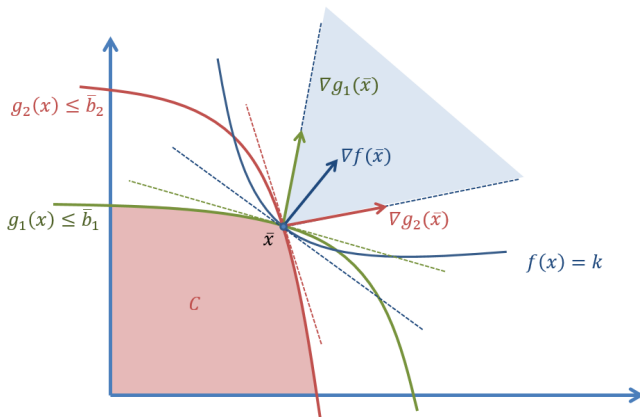
$$\lambda_m(g_m(\bar{x}) - \bar{b}_m) = 0$$

*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  $g_m(\bar{x}) \leq \bar{b}_m$  is binding.

# Kuhn-Tucker Conditions

## Explanation of Kuhn-Tucker Condition (i)





## Kuhn-Tucker Conditions

### Explanation of Kuhn-Tucker Condition (ii)

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

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

## 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  $\bar{x}$  under consideration.

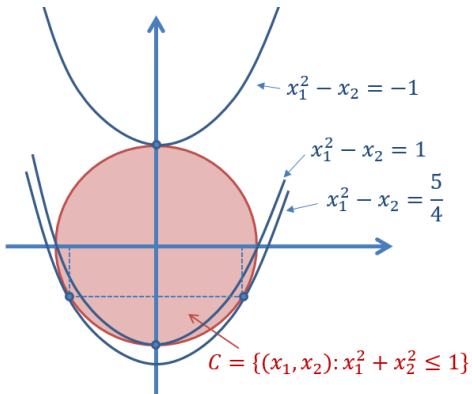
## Example of Inequality Constraints

**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

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

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

$$\vdots$$

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

$$x_1 \geq 0$$

$$\vdots$$

$$x_N \geq 0$$

## Adding Non-negativity Constraints

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

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