

# KKT conditions

Sunday, September 3, 2017 17:24

For the glory of God

## Necessary Condition for Constrained optimization

If we can remember correctly,  $\nabla f(x^*) = 0$  is necessary condition for unconstrained optimization. (1st order)

Then, Can we use the condition for constrained optimization?

The answer is No

↳ For constraint optimization, the necessary condition must account for the constraints.

⇒ Karush Kuhn Tucker (KKT) conditions are introduced. (This is 1st order necessary condition for constrained optimization)

; Here, please note that

- Even if you can solve a constrained optimization problem using the KKT conditions to find a point  $x^*$ ,

we are not guaranteed that the point is a local optimum.

- This is because the KKT conditions are just necessary condition.

- It could be a sufficient condition only if the function is convex.

what if we need to say 2nd order?

: Then, let  $L = \nabla L$ , i.e.  $\nabla(\nabla L) = 0$

for 2nd condition

The solution of a constrained optimization problem can often be found by using the so-called Lagrangian function.

## Lagrangian function ( $L$ )

Analysis of constrained optimization is based upon the Lagrangian function.

The Lagrangian function modifies the objective function to account for both inequality and equality constraints.

$$L(x, \lambda) = f(x) + \sum_{j=1}^m \lambda_j g_j(x) + \sum_{k=1}^l \lambda_{m+k} h_k(x) ; \text{ where } f(x) = \text{objective function}$$

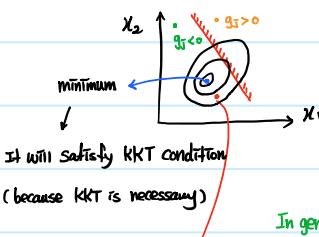
$\lambda_j = \lambda_{m+k}$  = Lagrange multipliers

Here,

-  $\lambda_{m+k}$  are unrestricted in sign because it doesn't matter for Equality constraint. ( $\because$  on the line)

-  $\lambda_j$  should be greater than zero because it does matter for inequality constraint. ( $\because$  one-side allowed, the other side non-allowed)

↳ We strict the sign of  $\lambda_j$  to make sure correct side for optimum.



Then, why  $\lambda_j \geq 0$ ? For inequality constraint, we know that  $\lambda_j \neq 0$  based on the complementary slackness.

: The condition  $\lambda_j \geq 0$  is based on the standard form.

If we define  $g_j \geq 0$  in the standard form, we would find  $\lambda_j \leq 0$ .

↳ Then,  $\lambda_j$  = positive or negative?

: According to the def. of standard form,

we know that  $g_j(x) \leq 0$  for inequality

if  $g_j = 0$ ,  $\Rightarrow$  if  $\lambda_j < 0$ ,  $\lambda_j g_j > 0$ ;  $\lambda = \text{const.}$

then if  $g_j < 0$

$\lambda_j > 0$ ,  $\lambda_j g_j < 0$

$\lambda_j = 0$ ,  $\lambda_j g_j = 0$  (but  $g_j \leq 0$ )

$g_j \leq 0$  still satisfy the form

We can check KKT at this point:

minimize  $f(x)$

however, I don't think that

KKT conditions are satisfied. Subject to  $g(x) \leq 0$  : Inequality Constraints  
(It's still in feasible region though)

$h(x) = 0$  : Equality Constraints

$x_L \leq x \leq x_U$  : Side constraints

This is so-called Standard form  
of constrained optimization,

More restricted..

Let  $g_j \leq 0$

(Satisfy the form)

Standard

of. The most worst optimization problem is the case where a lot of Equality constraints exist. ( $\because$  We have to be on that line)

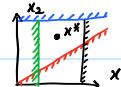
KKT Conditions

In terms of the conditions that we will discuss, please make sure that we need to satisfy all of them.

a) Check if  $x^*$  is feasible or not with respect to not only equality but inequality constraints

- If infeasible, stop it!

$x^*$  is said to be feasible if it satisfies all of constraints



b) Check  $\lambda_j g_j(x^*) = 0$  ;  $j = 1, 2, 3, \dots, m$

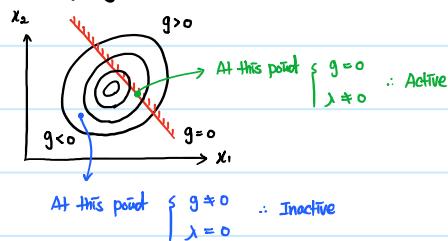
- This is called as Complementary Slackness.

- There are two possibilities to satisfy it

$\lambda_j = 0$ and $g_j \neq 0$	$\lambda_j \neq 0$ and $g_j = 0$
----------------------------------	----------------------------------

what if  $\lambda_j = 0$  and  $g_j = 0$  so that  $\lambda_j g_j = 0$ ?

< Inequality Constraint >



In short, that is not possible.

Since the inequality constraint can either be inactive or active but not both at the same time.

Therefore,  $\lambda_j$  and  $g_j$  cannot be 0 at the same time.

An inequality constraint is said to be inactive (or passive) at a potential solution  $x^*$

If it satisfies  $g(x^*) < 0$

c) Check  $\nabla_x L(x^*, \lambda) = 0$  for  $\lambda_j \geq 0$  For 2<sup>nd</sup> condition,  $\nabla(\nabla f) = 0$

Again, defined from the Standard form

The derivative of the equation will give us a solution, then we need to check if it satisfies  $\lambda_j \geq 0$

$$L(x, \lambda) = f(x) + \sum_{j=1}^m \lambda_j g_j(x) + \sum_{k=1}^l \lambda_{m+k} h_k(x)$$

$$\Leftrightarrow \nabla_x L(x^*, \lambda) = \underbrace{\nabla_x f(x^*)}_{A} + \underbrace{\sum_{j=1}^m \lambda_j \nabla_x g_j(x^*)}_{B}; \text{ where we assume that }$$

No Equality constraint

Given point, it would be an optimum if KKT are satisfied

Then,

A : This is a vector, pointing in the steepest direction of the objective function.

B : This is a linear combination of vectors. The set of all such vectors is called as a convex cone.

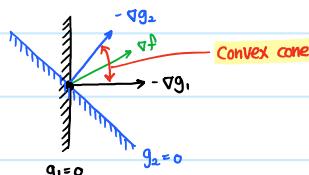
-  $\nabla_x L(x^*, \lambda) = 0$  means :

$$\Leftrightarrow A+B=0 \therefore B=-A$$

; In other words, the negative of the gradient of the objective function must be equal to some vector in the convex cone.

(or the negative of the gradient of the objective function must be in the convex cone)

- For instance, let's say ; we have two constraints



from the standard form, we already defined  $g_j \leq 0$

; where the vector length would be determined by  $\lambda$

constraints aren't linearly independent

What if  $\nabla f$  is outside of the convex cone?

: It wouldn't improve. (sometimes, it would violate some constraints.)

Then, Do the KKT condition always apply?  $\rightarrow$  In the same way,  $x^*$  is optimum  $\Leftrightarrow$  KKT condition is satisfied

o Always except for the case

• It wouldn't improve. (sometimes, it would violate some constraints)

Then, Do the KKT condition always apply?  $\rightarrow$  In the same way.  $x^*$  is optimum  $\Leftrightarrow$  KKT condition is satisfied

• The answer is No

• The KKT conditions apply only if the constraints satisfy some regularity condition, called **constraint qualification**.

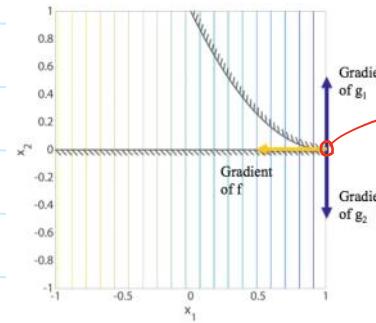
↑  
not always except for the convex case  
↓

The gradients of the active constraints must be linearly independent at the point at which the KKT conditions are being checked.

e.g. minimize  $f = -x_1$

Subject to  $x_2 \geq 0$

$$x_2 + (x_1 - 1)^3 \leq 0$$



- At this point,
- Not independent on  $g_1$  and  $g_2$
  - Can't check KKT conditions

↳ It will break 3) condition:

$$\nabla_x \mathcal{L} = 0$$

$$\Leftrightarrow [\nabla g_1, \nabla g_2, \dots] \lambda = -\nabla f$$

If  $g_1, g_2$  and  $g_3$  are dependent,  $g_2$  can be expressed by  $g_1 + g_2 \Rightarrow$  can't get  $\lambda$

#### Example application of the KKT conditions

• This is from AE6310 hand-written note :

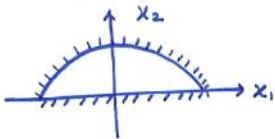
• Example application of the KKT conditions.

$$\min. f(x_1, x_2) = x_1 + x_2 *$$

$$\text{Subject to. } g_1(x_1, x_2) = x_1^2 + x_2^2 - 1 \leq 0$$

$$g_2(x_1, x_2) = -x_2 \leq 0$$

1) Check feasible region



2) Check  $\lambda_1 g_1(x^*) = 0$

$$L(x, \lambda) = (x_1 + x_2) + \lambda_1 (x_1^2 + x_2^2 - 1) + \lambda_2 (-x_2)$$

$$\text{Then, } \lambda_1 (x_1^2 + x_2^2 - 1) = 0$$

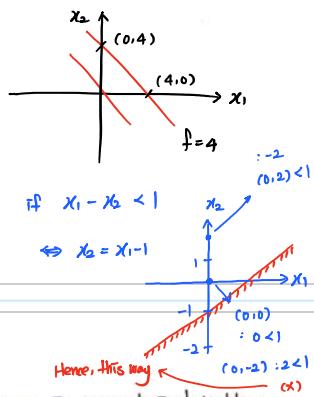
$$\lambda_2 (-x_2) = 0$$

3) Check  $\nabla_x L = 0$

$$\frac{\partial L}{\partial x_1} = 1 + 2\lambda_1 x_1 = 0$$

$$\frac{\partial L}{\partial x_2} = 1 + 2\lambda_1 x_2 - \lambda_2 = 0$$

$$\text{of. } f(x_1, x_2) = x_1 + x_2$$



4) Finally, we get

If the problem is unconstrained problem.

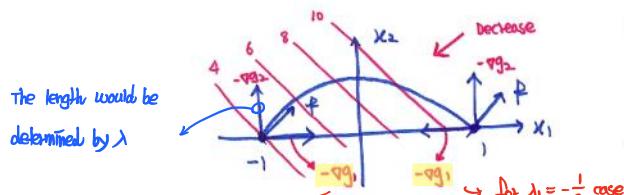
(keep decrease)

$$x_1 = \pm 1, x_2 = 0, \lambda_1 = -\frac{1}{2x_1}, \text{ and } \lambda_2 = 1$$

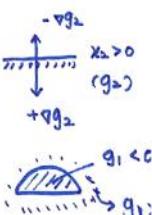
: There are two points that satisfy KKT conditions 2 and 3.

Which could be the constrained minimum?

The length would be determined by  $\lambda$



for  $\lambda_1 = \frac{1}{2}$  case Mathematically,  $x_1 = -1$  ( $\because \lambda_1 = \frac{1}{2} \geq 0$ ) of course, when  $x_2 = 0$



for  $\lambda_1 = \frac{1}{2}$  case Mathematically,  $x_1 = -1$  ( $\because \lambda_1 = \frac{1}{2} \geq 0$ ) of course, when  $x_2 = 0$   
 Graphically,  $x_1 = -1$  ( $\because$  vector is in convex) + just  $\frac{1}{2}$   
 at glance.  
 Since  $g_3 \leq 0$   
 for  $\lambda_3 \geq 0$ ,  
 Then, what is convex and so on? This is gonna be  
 the direction  
 would go inside of talking about the implications of  $\nabla f \perp L(x^*, \lambda) = 0$ .  
 the feasible region  
 In order to explain it, we need to go back to  
 the 3rd KKT condition.

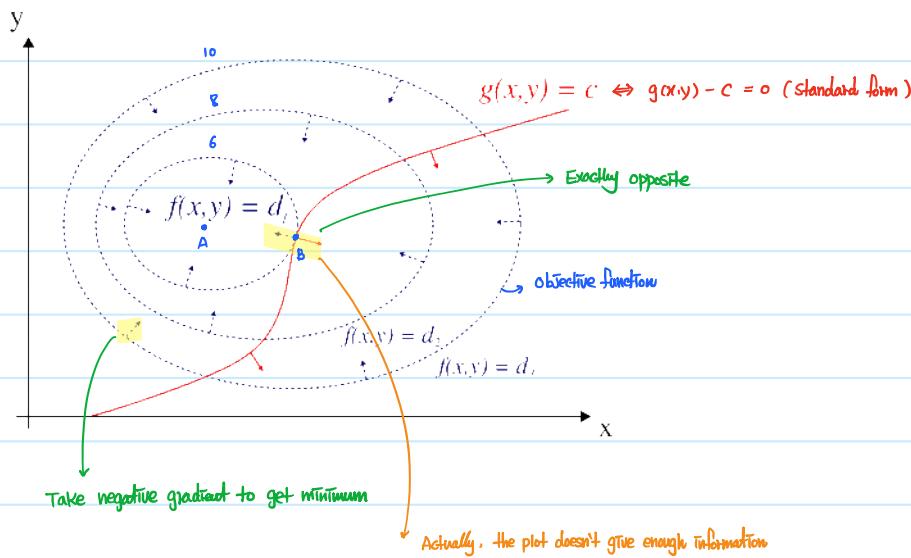
→ Next page

As a final note, what if there is only one active inequality constraint?

The convex cone collapses to a line, so the objective function's gradient must point in exactly the opposite direction as the inequality constraint's gradient.

$$\begin{aligned}\nabla f(x^*, \lambda) &= \nabla f(x^*) + \lambda \nabla g(x^*) = 0 \\ \Leftrightarrow \lambda \nabla g(x^*) &= -\nabla f(x^*)\end{aligned}$$

It implies that  $\nabla f$  and  $\nabla g$  lie on the same line

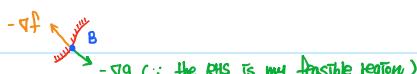


let's think about two cases:

1) : For this case, a minimum should be located around point A

same direction for  $-\nabla f$  and  $-\nabla g$

2) : For this case, a minimum should be located around point B



: this means  $\nabla g \parallel \nabla f$  on the same line

  $-\nabla g$  ( $\because$  the RHS is my feasible region) ; this means  $\nabla g \parallel \nabla f$  on the same line