

MATH - CAMP RECITATION FIVE

APPLYING THE

KARUSH-KUHN-TUCKER THEOREM

THE THEOREM

Let X be an open set in \mathbb{R}^n (generally \mathbb{R}^n)

$$f: X \rightarrow \mathbb{R} \quad C^1$$

$$\forall j \in \{1, \dots, k\}, g_j: X \rightarrow \mathbb{R} \quad C^1$$

$$\forall l \in \{1, \dots, m\}, h_l: X \rightarrow \mathbb{R} \quad C^1$$

If $\underset{\sim}{x}^* \in \mathbb{R}^n$ is a solution to

$$\max_{x \in X} f(x)$$

$$\text{s.t. } g_j(x) \geq 0 \quad \forall j$$

$$h_l(x) = 0 \quad \forall l$$

(and constraint qualification holds at x^*)

THEN

$$\exists \lambda_j, \mu_l \quad \forall j, l$$

$$\text{s.t. } \nabla f(x^*) + \lambda^T g'(x^*) + \mu^T h'(x^*) = \underset{\sim}{0}$$

That is, the system

$$\frac{\partial f(x)}{\partial x_1} + \sum_{j=1}^k \lambda_j \frac{\partial g_j(x)}{\partial x_1} + \sum_{l=1}^m \frac{\partial h_l(x)}{\partial x_1} = 0$$

⋮

$$\frac{\partial f(x)}{\partial x_n} + \sum_{j=1}^k \lambda_j \frac{\partial g_j(x)}{\partial x_n} + \sum_{l=1}^m \frac{\partial h_l(x)}{\partial x_n} = 0$$

has a solution at x^*

AND

Complementary Slackness,

which includes two conditions:

i) All h constraints bind.

$$\text{i.e. } h_l(x^*) = 0 \quad \forall l$$

ii) All g constraints either bind at x^* or $\lambda_j = 0$

$$(\lambda_j = 0) \wedge (g_j(x^*) > 0)$$

$$\text{OR } (\lambda_j > 0) \wedge (g_j(x^*) = 0)$$

INTERPRETATION OF APPLYING KKT

So, this is a x^* maximizer \Rightarrow

conditions are met at x^*

We can solve for when conditions are met. Then we have

Conditions are met at x^*

$\Rightarrow x^*$ is a POSSIBLE MAXIMIZER

So, we can think of KKT as a METHOD OF OBTAINING POTENTIAL

MAXIMA. (CQ gets in the way)

METHOD OF APPLYING KKT

STEP ONE: DOES THE PROBLEM HAVE A SOLUTION?

• To prove this, we must apply the Weierstrass theorem

Substep one: Show that the constraint set

$$D = \{x \in \mathbb{R}^n : g_j(x) \geq 0, h_l(x) = 0 \forall j, l\}$$

is closed and bounded (therefore compact by Heine-Borel)

Substep two: Show that $f(x)$ is continuous over D

Substep three: "apply Weierstrass"

STEP TWO:

IMPROVE THAT SOME INEQUALITY CONSTRAINTS BIND OR DO NOT BIND IN

MAXIMUM - Inequality constraints make things difficult because of the complementary slackness conditions. So, we want to either prove that

$g_j(x^*) = 0$ or $g_j(x^*) > 0 \Rightarrow \lambda_j = 0$
for as many as possible before we solve Lagrangian for possible maximum candidates.

THE TWO MOST POPULAR WAYS TO DO SO IS:

1) Binding by strict monotonicity:

If the objective function is strictly increasing in a set of x 's (\tilde{x}), and $\tilde{g}(\tilde{x})$, an inequality constraint, is strictly increasing in that set of x 's, and no other constraint gets in the way of increasing x 's in that set, THEN $\tilde{g}(\tilde{x})$ will bind at the maximum.

We generally prove on a case by case basis. Example:

$$\max M^\alpha C^\beta \quad 0 < \alpha, \beta < 1$$

$$\text{s.t. } (80)M + (60)C \leq 300$$

Proof that constraint will bind:

Assume otherwise. $\exists S^0, C^0$ that are maxima
and $80S^0 + 60C^0 < 300$

$$\Rightarrow \exists \epsilon > 0 \text{ s.t. } 80S^0 + 60C^0 + \epsilon < 300$$

$$\Rightarrow \exists 80S^0 + 60\left(C^0 + \frac{\epsilon}{60}\right) < 300$$

$$\left(80S^0\right)^\alpha \left(60\left(C^0 + \frac{\epsilon}{60}\right)\right)^\beta > \left(80S^0\right)^\alpha \left(60C^0\right)^\beta$$

which contradicts the statement that S^0, C^0 is a maximum.

2) Not BINDING BY CONTRADICTION

To prove a constraint is not binding at the maximum, assume the constraint is binding at a maximum, then find a point in D that strictly dominates the binding points but where the constraint is not binding. This contradicts the statement that it is a max.

Example:

$$\max M^\alpha C^\beta$$

$$\text{s.t. } 80M + 60C \leq 300$$

$$C \geq 0$$

$$M \geq 0$$

Assume C^0, M^0 is a max and $C^0 = 0$

$$M^0 \alpha C^0 \beta = 0$$

But $(C^0, M^0) = (1, 1)$ satisfies all constraints:

$$\bullet \quad 80 + 60 \leq 300$$

$$1 \geq 0$$

$$1 \geq 0$$

and dominates C^0, M^0 :

$$1^\alpha 1^\beta = 1 > 0$$

So M^0, C^0 is not a maximum.

STEP THREE:

SOLVE THE LAGRANGIAN PROBLEM
FOR ALL REMAINING COMBINATIONS
OF COMPLEMENTARY SLACKNESS
CONDITIONS

We started off with the problem

$$\begin{aligned} \max_{x \in X} & f(x) \\ \text{s.t.} & g_j(x) \geq 0 \quad j \in \{1, \dots, k\} \\ & h_l(x) = 0 \quad l \in \{1, \dots, m\} \end{aligned}$$

In step 2, we narrowed this
down to

$$\begin{aligned} \max_{x \in X} & f(x) \\ \text{s.t.} & g_r(x) \geq 0 \quad \forall r \text{ s.t. we still don't know} \\ & g_k(x) = 0 \quad \forall k \text{ s.t. we proved it binds} \\ & h_l(x) = 0 \end{aligned}$$

Now, for every r , we have 2
options: BIND / NOT BIND.

Say $r \in \{1, \dots, R\}$

Then, there are 2^R different Lagrangian problems we must solve:

$$\left\{ \begin{array}{l} \max_{x \in X} f(x) \\ \text{s.t. } g_{r'}(x) = 0 \quad \forall r' \in S \\ g_n(x) = 0 \quad \forall n \\ h_u(x) = 0 \end{array} \right. : S \in \mathcal{P}(\{1, \dots, R\})$$

↓
Power set

↓
Set of problems to apply Lagrangian method to.

So, for each of these problems, write out the system

$$\nabla f(x) + \lambda^T [g_{r'}'(x), g_n'(x)] + \mu^T h'(x) = \underset{0}{\mathbf{0}}$$

and solve for solutions.

STEP FOUR:

SEARCH FOR ALL POINTS AT WHICH THE CONSTRAINT QUALIFICATION DOES NOT HOLD, THAT IS, FOR EACH OF THE 2^R PROBLEMS DEFINED BY $S \subseteq P(\{1, \dots, R\})$ IN STEP 3, FIND ALL POINTS WHERE

$\nabla g_0(x) \cup \nabla g_n(x) \cup \nabla h_e(x)$
are linearly dependent.

STEP FIVE: EVALUATE $f(x)$ AT ALL THE POINTS ATTAINED IN STEPS 3 & 4. THE LARGEST IS (ARE) THE MAXIMUM (MAXIMA), UNDER STEP ONE.

ALTERNATIVE METHOD:

USING SUFFICIENT CONDITIONS
AS LAID OUT IN THEOREM 4.5 or Other.

IN THIS METHOD, FOLLOWS STEPS
2-3, THEN FOR EACH POTENTIAL
CASE, CHECK IF THE SUFFICIENT
CONDITIONS HOLD. IF THEY HOLD AT x^* ,
 x^* IS A (NOT NECESSARILY UNIQUE) MAXIMIZER.

Example: under regularity
conditions, sufficient conditions are:

$$0 \quad \begin{bmatrix} \frac{\partial^2 \mathcal{L}(x^*, \lambda^*)}{\partial x_1 \partial x_1} & \dots & \frac{\partial^2 \mathcal{L}(x^*, \lambda^*)}{\partial x_n \partial x_1} \\ \vdots & & \vdots \\ \frac{\partial^2 \mathcal{L}(x^*, \lambda^*)}{\partial x_1 \partial x_n} & \dots & \frac{\partial^2 \mathcal{L}(x^*, \lambda^*)}{\partial x_n \partial x_n} \end{bmatrix} = \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) \text{ is} \\ \text{positive} \\ \text{semi-definite}$$

OR, equivalently, $\mathcal{L}(x, \lambda)$ is concave in x at (x^*, λ^*)

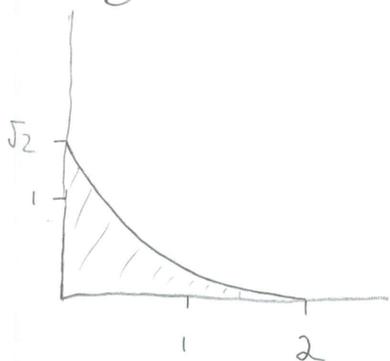
IF AT NO MAXIMIZER IS FOUND, USE
THE FIRST METHOD.

EXAMPLES

$$\begin{aligned} \max_{x, y \in \mathbb{R}} \quad & xy \\ \text{s.t.} \quad & x + y^2 \leq 2 \\ & x \geq 0 \\ & y \geq 0 \end{aligned}$$

STEP ONE:

Very fast:



Take any (x, y) on $x + y^2 = 2$ boundary.

$$\text{If } x=0, \exists \epsilon \text{ s.t. } x + (y-\epsilon)^2 \leq 2 \quad x=0, y \geq 0$$

$$\text{If } y=0, \exists \epsilon \text{ s.t. } (x-\epsilon) + y^2 \leq 2 \quad x \geq 0, y=0$$

The x & y boundary is trivial

so D is closed.

$B_3(0,0)$ bounds D

$\Rightarrow D$ compact in \mathbb{R}^2 by H-B

xy is continuous over \mathbb{R}^2

\Rightarrow maximum exists by Weierstrass.

12/02/2021

STEP 2:

$$x \geq 0$$

$y \geq 0$ do not bind, as

if $x=0$, $\max_{x,y \in \mathbb{R}} xy = 0 \quad \forall y$,

but $(x,y) = (1,1)$ is feasible

$$\text{and } xy|_{(1,1)} = 1 > 0$$

$x + y^2 \leq 2$ does bind, as

xy is strictly increasing in x & y

and $g = 2 - x - y^2$ is strictly decreasing in x & y .

STEP 3: Now, we only have one possible problem to apply Lagrangian

$$\text{TO: } \max_{x,y \in \mathbb{R}} xy$$

$$\text{s.t. } x + y^2 = 2$$

$$\mathcal{L}(x,y) = xy + \lambda(2 - x + y^2)$$

FOCs:

$$\frac{\partial f}{\partial x} = y - \lambda = 0 \quad (1)$$

$$\frac{\partial f}{\partial y} = x - 2\lambda y = 0 \quad (2)$$

$$\frac{\partial f}{\partial \lambda} = 2 - x - y^2 = 0 \quad (3)$$

Plugging (1) into (2):

$$x = 2y^2$$

into (3):

$$2 = 2y^2 + y^2$$

$$\Rightarrow y = \pm \sqrt{2/3}$$

only $+\sqrt{2/3}$ is in D .

$$\text{so } y^* = \sqrt{2/3}$$

$$x^* = 4/3$$

$$\text{Step 4: } \nabla (2 - x - y^2) = \begin{bmatrix} -1 \\ -2y \end{bmatrix},$$

is a linearly independent set of vectors for all x, y .

So $(4/3, \sqrt{2/3})$ is the only maximizer.

EXAMPLE 2:

$$\max_{x_1, x_2 \in \mathbb{R}} x_1$$

$$\text{s.t. } (1-x_1)^3 - x_2 \geq 0$$

$$x_1 \geq 0$$

$$x_2 \geq 0$$

STEP 1: Similar argument as always.

STEP 2: $x_1 \geq 0$ will not bind, as

if $x_1 = 0$ $f(x_2) = 0$ and $(x_1, x_2) = (1, 0)$ is feasible and $f((1, 0)) = 1$

$(1-x_1)^3 - x_2 \geq 0$ will bind, as

x_1 is strictly increasing in x_1 ,

$(1-x_1)^3 - x_2$ is strictly decreasing in x_1

and increasing x_1 does not run into any

other constraints.

$x_2 \geq 0$ will always bind as if $x_2^0 > 0$,

$\exists \epsilon \text{ s.t. } (x_1^0 + \epsilon) > x_1^0 \text{ and } (x_2^0 - \epsilon) < x_2^0$.

STEP 3: We have two possible

Lagrangian problems now:

~~$$\max_{x_1, x_2 \in \mathbb{R}} x_1$$~~

~~$$\text{s.t. } (1-x_1)^3 - x_2 = 0$$~~

~~$$x_2 = 0$$~~

~~$$\max_{x_1, x_2 \in \mathbb{R}} x_1$$~~
~~$$\text{s.t. } (1-x_1)^3 + x_2 = 0$$~~

The second

$$f = x_1 + \lambda((1-x_1)^3 - x_2)$$

FOCs

$$\frac{\partial f}{\partial x_1} = 1 - \lambda \cdot 3(1-x_1)^2 = 0$$

$$\frac{\partial f}{\partial x_2} = -\lambda = 0$$

$$\frac{\partial f}{\partial \lambda} = (1-x_1)^3 - x_2 = 0$$

This does not tell us anything!
All points that satisfy the constraint
are included in our set of potential
maximizers.

The first

$$f = x_1 + \lambda_1((1-x_1)^3 - x_2) + \lambda_2(x_2)$$

$$\frac{\partial f}{\partial x_1} = 1 - \lambda_1 \cdot 3(1-x_1)^2 = 0 \quad (1)$$

$$\frac{\partial f}{\partial x_2} = -\lambda_1 + \lambda_2 = 0 \quad (2)$$

$$\frac{\partial f}{\partial \lambda_1} = (1-x_1)^3 - x_2 = 0 \quad (3)$$

$$\frac{\partial f}{\partial \lambda_2} = x_2 = 0 \quad (4)$$

$$(4) \& (3) \Rightarrow (1-x_1)^3 = 0$$

$$\Rightarrow x_1 = 1$$

But! Note that we can't find a λ_1, λ_2 , so we cannot officially apply KKT

↳ Usually, when this happens, the following is the case.

STEP FOUR: first ~~then second~~

$$\nabla [(1-x_1)^3 - x_2] = \begin{bmatrix} -3(1-x_1)^2 \\ -1 \end{bmatrix},$$

which is always linearly dependent

~~then second~~

~~$\nabla [(1-x_1)^3 - x_2]$ as above~~

$$\nabla [x_2] = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -3(1-x_1)^2 \\ -1 \end{bmatrix} \& \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

are linearly dependent when $x_1 = 1$, so we need to check all of these points too. Luckily, only $(1,0) \in D$ where $x_1 = 1$! 16

~~STEP FIVE: CHECK ALL POSSIBLE VALUES.~~

~~From the second:~~

STEP FIVE: CHECK ALL POSSIBLE
VALUES

Only one value, so it
must be the solution!

$$\begin{array}{l} x_1^* = 1 \\ x_2^* = 0 \end{array}$$