

Constrained optimization and KKT Conditions

Understanding KKT conditions

AUTHORS
Rahul Vigneswaran

PUBLISHED
Feb. 8, 2024

AFFILIATIONS
IITH

Contents

[Necessity Vs Sufficiency](#)

[Primal and Dual](#)

[KKT Conditions](#)

[Necessity](#)

[Sufficiency](#)

[Conclusion](#)

Necessity Vs Sufficiency

Proving something requires proving both the necessity and sufficiency. So lets first understand what they mean.

Lets take A and B .

Necessity: $A \implies B$

- Assume A is true
- Now show B is true as a consequence of the above assumption.
- When we say A is necessary for B , we mean that for B to be true, A must be true.

Sufficiency: $A \longleftarrow B$

- Assume B is true.
- Now show A is true as a consequence of B .
- When we say A is sufficient for B , we mean that for A to be true, B must be true.

Necessary and sufficient: $A \iff B$

- If A is true, B is definitely true.
- If B is true, A is definitely true.

For more check, [Source 1](#).

Primal and Dual

Lets first establish some common notations before we prove anything The unconstrained optimisation problem is,

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & h_{i(x)} \leq 0, \forall i \quad [\text{eqn}(4)] \\ & l_{j(x)} = 0, \forall j \quad [\text{eqn}(3)] \end{aligned}$$

This above is our primal.

The dual for it is,

$$\begin{aligned} \max_{u,v} \min_x \quad & f(x) + \sum_{i=1}^m u_i h_i(x) + \sum_{j=1}^r v_j l_j(x) \\ \text{s.t.} \quad & u \geq 0 \quad [\text{eqn}(5)] \end{aligned}$$

For the sake of simplicity lets define,

$$\text{Lagrangian: } L(x, u, v) = f(x) + \sum_{i=1}^m u_i h_i(x) + \sum_{j=1}^r v_j l_j(x) \quad [\text{eqn}(1)]$$

$$g(u, v) = \min_x L(x, u, v) \quad [\text{eqn}(2)]$$

Dual: $\max_{u,v} g(u, v)$

Solution of primal : x^*

Solution of dual : u^*, v^*

KKT Conditions

The KKT conditions are as follows,

1. Stationary condition

$$0 \in \partial \left(f(x) + \sum_{i=1}^m u_i h_i(x) + \sum_{j=1}^r v_j l_j(x) \right)$$

1. Complementary slackness

$$u_i \cdot h_i(x) = 0, \quad \forall i \in 1, \dots, m$$

Why is there no complementary slackness for $v_j l_j$? Because it is a equality constraint. At optimal x it will always be 0.

2. Primal feasibility

$$\begin{aligned} h_{i(x)} &\leq 0, \forall i \in 1, \dots, m \\ l_{j(x)} &= 0, \forall j \in 1, \dots, r \end{aligned}$$

3. Dual feasibility

$$u_{i(x)} \geq 0, \forall i \in 1, \dots, m$$

Why is there no dual feasibility for v_j ? I dont know the answer yet. Will update once I understand it.

We are now going to prove **Optimality** \iff **KKT**. To do that, like we saw earlier we need to prove both necessity and sufficiency.

Necessity

$$\text{Optimality} \implies \text{KKT}$$

We are also assuming strong duality (Look into Slater's condition) and as a result there is no duality gap, that is $f(x^*) - g(u^*, v^*) = 0$

Lets first not worry about the KKT at all and solve as if we are solving a generic equation.

$$\begin{aligned} f(x^*) &= g(u^*, v^*) \\ &= \min_x f(x) + \sum_{i=1}^m u_i^* h_i(x) + \sum_{j=1}^r v_j^* l_j(x) \\ &\leq f(x^*) + \sum_{i=1}^m u_i^* h_i(x^*) + \sum_{j=1}^r v_j^* l_j(x^*) \\ &\leq f(x^*) \end{aligned}$$

Lets go through the above now step by step.

- Step 1: $f(x^*) = g(u^*, v^*)$
 - This is true because there is no duality gap.
- Step 2: $= \min_x f(x) + \sum_{i=1}^m u_i^* h_i(x) + \sum_{j=1}^r v_j^* l_j(x)$
 - The above is by eqn (1) and eqn (2). We have already defined it. There is nothing special here.
- Step 3: $\leq f(x^*) + \sum_{i=1}^m u_i^* h_i(x^*) + \sum_{j=1}^r v_j^* l_j(x^*)$
 - Lets understand how we went from step 2 to step 3. x^* is a solution for the primal but for g it is just some feasible solution. There is no guarantee that it is an optimal. Therefore it must be greater than equal to the actual optimal solution of $f(x) + \sum_{i=1}^m u_i^* h_i(x) + \sum_{j=1}^r v_j^* l_j(x)$.
- Step 4: $\leq f(x^*)$
 - Lets understand how we went from step 3 to step 4. Its is two parted.
 - Lets first focus on $\sum_{j=1}^r v_j^* l_j(x)$.
 - We already know that from eqn (3) that $v_j^* l_j(x^*)$ must be 0 if x^* is the optimal if the constraints are respected. So summation of all individual j -s will also be 0. Therefore the term entirely will be zero.
 - Now lets focus on $\sum_{i=1}^m u_i^* h_i(x)$
 - From eqn (4) we know that $h_i(x)$ must be -ve. Then from eqn (5) we know that u_i^* must be +ve. As a result $u_i^* h_i(x)$ is +ve multiplied by -ve, should be -ve too. When we sum up all i which are individually negative values, the resulting summation would be -ve too.
 - Now lets regroup. (something) + (-ve value) + (zero) \leq (something). Therefore, $f(x^*) + \sum_{i=1}^m u_i^* h_i(x^*) + \sum_{j=1}^r v_j^* l_j(x^*) \leq f(x^*)$.

- \leq in step 4
 - Lets first handle the inequality in step 4 as it is much easier to understand.
 - We already saw how we transitioned from step 3 to step 4. Now lets see how we can make the inequality into an equality.

$$f(x^*) + \sum_{i=1}^m u_i^* h_i(x^*) + \sum_{j=1}^r v_j^* l_j(x^*) \leq f(x^*)$$

- $\sum_{j=1}^r v_j^* l_j(x^*)$:
 - We have already shown that this term is 0. So we dont have to worry about it.
- $\sum_{i=1}^m u_i^* h_i(x^*)$:
 - Now we have to show that this term is 0 in order for the inequality to become an equality. But we saw earlier that this term is actually -ve and not 0. So now comes the KKT condition 2 - The complimentary slackness condition. This enforces that $u_i \cdot h_i(x) = 0$. Since each term is 0, the summation entirely is 0. This makes the inequality into an equality. Without the KKT condition 2 this would have not been possible.

- \leq in step 3
 - Now lets handle the only inequality left. This was slightly difficult for me to come to terms with.

$$\min_x f(x) + \sum_{i=1}^m u_i^* h_i(x) + \sum_{j=1}^r v_j^* l_j(x) \leq f(x^*) + \sum_{i=1}^m u_i^* h_i(x^*) + \sum_{j=1}^r v_j^* l_j(x^*)$$

- We already saw that x^* is actually not an optimal solution for the LHS (that is $g(x)$) but is just some feasible solution. So for the inequality to become an equality we have to somehow enforce that x^* is also an optimal solution for $g(x)$.
 - Lets focus only on the LHS : $\min_x f(x) + \sum_{i=1}^m u_i^* h_i(x) + \sum_{j=1}^r v_j^* l_j(x)$
 - Note that from eqn (1), $L(x, u^*, v^*) = f(x) + \sum_{i=1}^m u_i^* h_i(x) + \sum_{j=1}^r v_j^* l_j(x)$. So the LHS can also be written as $\min_x L(x, u^*, v^*)$.
 - Say the optimal solution for this after minimising wrt x is \tilde{x} . Note that as of now $\tilde{x} \neq x^*$. That is what we have to enforce but we will get there soon. For now remember that \tilde{x} is different and is the optimal solution for the above.
 - If \tilde{x} is the optimal solution then if we get the sub-gradient of L at \tilde{x} , that is $\partial L(\tilde{x}, u^*, v^*) = 0$. (Since we are using a convex differentiable function, we can consider that sub-gradient is the same as gradient.). In other words $0 \in \partial L(\tilde{x}, u^*, v^*)$.
 - Now to back to enforcing that $\tilde{x} = x^*$. Only if $\tilde{x} = x^*$, the inequality becomes an equality. This is where KKT condition 1 - Stationarity condition comes into play. It enforces this. If $\tilde{x} = x^*$ that means like we saw earlier, $0 \in \partial L(\tilde{x} = x^*, u^*, v^*)$ which is the same as $0 \in \partial \left(f(x^*) + \sum_{i=1}^m u_i^* h_i(x^*) + \sum_{j=1}^r v_j^* l_j(x^*) \right)$.

We have now shown that both the KKT condition 1 and 2 are required to prove necessity. The conditions 3 and 4 are by default required to enforce feasibility of the primal and the dual problem hence the condition names.

Hence necessity is proved.

Sufficiency

$$\text{Optimality} \Leftarrow \text{KKT}$$

Now lets try to prove the other direction. Compared to the necessity, the sufficiency here is much easier to prove and understand. So here we assume KKT conditions are true, now we need to prove optimality.

Say x^*, u^*, v^* satisfy the KKT conditions then we prove have to prove that they are the optimal values for primal and dual. That is we have to show

$$\begin{aligned}
g(u^*, v^*) &= \min_x f(x) + \sum_{i=1}^m u_i^* h_i(x) + \sum_{j=1}^r v_j^* l_j(x) \\
&= f(x^*) + \sum_{i=1}^m u_i^* h_i(x^*) + \sum_{j=1}^r v_j^* l_j(x^*) \\
&= f(x^*)
\end{aligned}$$

Lets go through the above now step by step,

- Step 1: $g(u^*, v^*) = \min_x f(x) + \sum_{i=1}^m u_i^* h_i(x) + \sum_{j=1}^r v_j^* l_j(x)$
 - We defined this in eqn (2)
- Step 2: $\min_x f(x) + \sum_{i=1}^m u_i^* h_i(x) + \sum_{j=1}^r v_j^* l_j(x) = f(x^*) + \sum_{i=1}^m u_i^* h_i(x^*) + \sum_{j=1}^r v_j^* l_j(x^*)$
 - This is true because the KKT condition 1 - the stationarity condition. Since from it we already know that $0 \in \partial \left(f(x^*) + \sum_{i=1}^m u_i^* h_i(x^*) + \sum_{j=1}^r v_j^* l_j(x^*) \right)$, x^* is the optimal solution for the \min_x .
- Step 3: $f(x^*) + \sum_{i=1}^m u_i^* h_i(x^*) + \sum_{j=1}^r v_j^* l_j(x^*) = f(x^*)$
 - This is true because of the KKT condition 2 - the complimentary slackness condition. Since from it we already know that $u_i^* h_i(x^*) = 0$, the summation of individual terms will also be zero.
 - By KKT condition 3 - primal feasibility condition, we already know that $l_j(x^*) = 0$, so the 3rd term becomes 0 too.

Hence sufficiency is proved.

Conclusion

We have now effectively proved both directions by proving necessity and sufficiency. To sum it up,

- For an constrained optimization problem, if x^*, u^*, v^* satisfy the KKT conditions, then we can say that satisfying those KKT conditions is sufficient enough to imply that x^*, u^*, v^* are the optimal primal and dual solutions for that problem.
- Also, if strong duality holds and there exist optimal solutions x^*, u^*, v^* for that problem, then they must necessarily satisfy the KKT conditions.

Acknowledgements

I heavily borrowed from [Ryan Tibshirani](#)'s video lecture and notes for understanding KKT conditions. There are other lectures of Ryan explaining KKT conditions but I personally felt this one was much more easier to understand.

Thank you [Piyushi](#) for helping me understand this topic and proof-reading this blog!

An error occurred: giscus is not installed on this repository