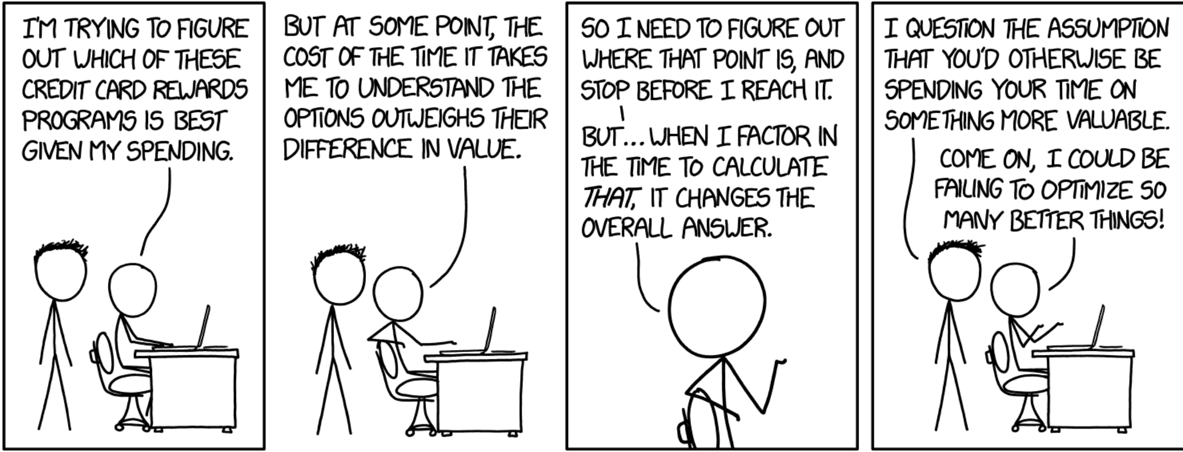
APEC Math Review

Khun Tucker

Natalia Ordaz Reynoso

Summer 2019

Optimization



Motivation

- We cannot simplify all problems to problems where we can use Lagrange's Theorem
- In particular, some inequality constraints cannot be replaced by an open set
- So we need another theorem: Khun Tucker
- We will see feasible sets of the form: $D = U \cap \{x \in \mathbb{R}^n : h_i(x) \geq 0, i = 1, ..., l\}$
- the Theorem of Khun Tucker describes necessary conditions for local optima in such problems.
- We say that the constraint $h_i(x) \geq 0$ is effective at a point x^* if the constraint holds with equality at x^*

Theorem of Khun and Tucker

- Let $f : \mathbb{R}^n \to \mathbb{R}$, and $h_i : \mathbb{R}^n \to \mathbb{R}$ be C^1 functions.
- Suppose x* is a local maximum of f on $D = U \cap \{x \in \mathbb{R}^n : h_i(x) \geq 0, \forall i = 1, ...l\}$, where U is an open set.
- Let $E \subset \{1, ..., l\}$ denote the set of effective constraints at x*.
- Let $h_E = (h_i)_{i \in E}$
- Suppose $\rho(Dh_E(x^*)) = |E|$

Then, there exists a vector $\lambda^* = (\lambda_1^*, ...\lambda_l^*) \in \mathbb{R}^l$ such that the following conditions are met:

1. $\lambda_i^* \geq 0$ and $\lambda_i^* h_i(x^*) = 0, \forall i = 1, ..., l$
2. $Df(x^*) + \sum_{i=1}^{L} \lambda_i Dh_i(x^*) = 0$

- The first condition is called the complimentary slackness (at most one inequality is slack)
- These are necessary conditions: not sufficient! Here is an example:

$f(x) = x^3$, $g(x) = x$. $x^* = \lambda^* = 0$ is not a maximum

The Constraint Qualification

- Note that we only require it for effective constraints at the maximum.
- Example of when the constraint qualification is not met: $f(x,y) = -(x^2 + y^2)$, $h(x,y) = (x-1)^3 - y^2$
- What does the set D look like?
- Inspect the function, where is it maximized?

- Note that we only require it for effective constraints at the maximum.
- Example of when the constraint qualification is not met: $f(x, y) = -(x^2 + y^2)$, $h(x, y) = (x - 1)^3 - y^2$
- What does the set D look like?
- Inspect the function, where is it maximized?
- $(x^*, y^*) = (1, 0)$
- The problem's single constraint is effective at this point
- What is $Dh(x^*, y^*)$
- The assumptions of the theorem fail, and so does its conclusion.

The Khun Tucker Multiplier

- KT multipliers also measure the sensitivity of the objective function at x^* to relaxations of its various constraints.
- This intuition is clear with the complementary slackness condition.

Khun Tucker Cookbook Procedure

For Maximization:

1. Set up the lagrangean
2. Find all solutions to the following set of equations:
    - $\frac{\delta L}{\delta x_j}(x, \lambda) = 0, \forall j = 1, ..., n$
    - $\frac{\delta L}{\delta \lambda_i}(x, \lambda) \geq 0,\ \lambda_i \frac{\delta L}{\delta \lambda_i}(x, \lambda) = 0,\ \lambda_i \geq 0, \forall i = 1, ...l$

    The points that solve those equations are called critical points of L, note that it is not the same set as with equality constraints
3. Compute the value of f at every critical point of L.

What about minimization??

$\rightarrow$ *Examples*

General Case: Mixed Constraints

- $D = U \cap \{x \in \mathbb{R}^n : g(x) = 0, h(x) \geq 0$
- U is open
- $g : \mathbb{R}^n \to \mathbb{R}^k$
- $h : \mathbb{R}^n \to \mathbb{R}^l$
- We can define ϕ_i to be the collection of all g and h functions.
- We make a lagrangean with all ϕ functions and apply Khun-Tucker.

6.6 Exercises

Convex Structures in Optimization Theory

- Why did we spend so much time with convexity?

Convex Structures in Optimization Theory

- Why did we spend so much time with convexity?
- In problems with a convex constraint set and a concave objective function, the FOC are both necessary and sufficient to identify global maxima

Convex Structures in Optimization Theory

- Why did we spend so much time with convexity?
- In problems with a convex constraint set and a concave objective function, the FOC are both necessary and sufficient to identify global maxima
- In problems with a convex constraint set and a convex objective function, the FOC are both necessary and sufficient to identify global minima
- Some additional regularity conditions need to be satisfied.

Convex Structures in Optimization Theory

- Why did we spend so much time with convexity?
- In problems with a convex constraint set and a concave objective function, the FOC are both necessary and sufficient to identify global maxima
- In problems with a convex constraint set and a convex objective function, the FOC are both necessary and sufficient to identify global minima
- Some additional regularity conditions need to be satisfied.

Furthermore...

- We will say that a maximization problem is a convex maximization problem if it has a convex constraint set and a concave objective function.
- A minimization problem is said to be convex if the constraint set is convex and the objective function is convex
  $\rightarrow$ "the optimization problem is convex"
- Strictly concave (convex) objective function in a convex maximization (minimization) problem also gives us uniqueness.
- This result is extremely useful in economics

Implications of convexity

1. Every concave or concave function has to be continuous on the interior of its domain
2. Every concave or concave function must have some differentiability properties (in particular, every directional derivative must be well-defined at all points in the domain)
   - A concave or convex function defined over an open set of $\mathbb{R}^n$ must be differentiable "almost anywhere" in its domain.
   - With respect to the Lebesgue measure
3. The concavity or convexity of an everywhere differentiable function can be completely characterized by the behavior of its derivative. Concavity or convexity of a C^2 function can be completely characterized by its second derivative

Point 3: Concavity or Convexity Characterized with Derivatives

- Theorem 7.9: Let D be an open and convex set in $\mathbb{R}^n$, and let $f : D \to \mathbb{R}$ be differentiable on D. Then, f is concave if and only if $Df(x)(y - x) \geq f(y) - f(x)$ for all $x, y \in D$
- Theorem 7.10: Let $f : D \to \mathbb{R}$ be a C^2 function. Then
  1. f is concave on D IFF $D^2 f(x)$ is NSD for all x in D
  2. f is convex on D IFF $D^2 f(x)$ is PSD for all x in D
  3. If $D^2 f(x)$ is ND for all x in D, f is strictly concave on D
  4. If $D^2 f(x)$ is PD for all x in D, f is strictly convex on D

3 and 4 are only sufficient, not necessary: for example $f(x) = x^{-4}$, $g(x) = x^4$
Example: $f(x, y) = x^a y^b$. Check concavity/convexity both ways.

- Theorem 7.13: Suppose $D \subset \mathbb{R}^n$ is convex and $f : D \to \mathbb{R}$ is concave. Then:
  1. Any local maximum of f is a global maximum of f
  2. The set $argmax\{f(x)|x \in D\}$ of maximizers of f over D is either empty or convex
- Theorem 7.14 $D \subset \mathbb{R}^n$ is convex and $f : D \to \mathbb{R}$ is strictly concave. Then $argmax\{f(x)|x \in D\}$ is either empty or it contains only one element.

What about convex functions? What about minimization?

- Theorem 7.15 $D \subset \mathbb{R}^n$ is convex and $f : D \to \mathbb{R}$ is concave and differentiable. Then x is an unconstrained maximum of f if and only if $Df(x) = 0$

Theorem 7.16 (KT under Convexity): Let f be a concave C^1 function mapping $U \subset \mathbb{R}^n$ onto the real numbers, where U is open and convex. For i=1,...,l let $h_i : U \to \mathbb{R}$ also be concave C^1 functions. Suppose $\exists \bar{x} \in U$ such that $h_i(\bar{x}) > 0, i = 1, ..., l$ (Slater's condition, for necessity, replaces the rank condition). Then, x^* maximizes f over $D = \{x \in U : h_i(x) \geq 0, i = 1, ..., l\}$ If and only if $\exists \lambda^* \in \mathbb{R}^k$ such that the Khun-Tucker FOC are satisfied:

- $Df(x*) + \sum_{i=1}^{l} \lambda_i^* Dh_i(x^*) = 0$
- $\lambda^* \geq 0$, $\sum_{i=1}^{l} \lambda_i^* h_i(x^*) = 0$

Example showing the relevance of Slater's condition: $f(x) = x$, $h(x) = -x^2$, $x \in \mathbb{R}$

Using Convexity in Optimization: Summary

max f(x) subject to $x \in D = \{zinU : h_i(z) \geq 0, i = 1, ..., l\}$

- functions f, h_i are C^1, concave on D, and U is open.
- $L(x, \lambda) = f(x) + \sum_{i=1}^{l} \lambda_i h_i(x)$
- The critical points of L are characterized by
    1. $Df(x) + \lambda_i Dh_i(x) = 0$
    2. $\lambda_i \geq 0, h_i(x) \geq 0, \lambda_i h_i(x) = 0$ i=1,...,l

Then, we have two cases:

Using Convexity in Optimization: if Slater's Condition Holds

- If Slater's condition is met, Theorem 7.16 we can use the Khun-Tucker cookbook straight away, without checking for the conditions of proposition 6.5 (optimum exists, and constraint qualification)
- The above is true because Slater's condition + concavity of f and h imply that the KT conditions are necessary at an optimum, but also because they imply that KT conditions are sufficient.
- $\Rightarrow$ If L has no critical points, no solution exists, and If $(*x^*, \lambda^*)$ is a critical point of L, x^* is a solution to the maximization problem.
- We don't have to compare the value of solutions: they are all global max

Using Convexity in Optimization: if Slater's Condition Does not Hold

- If Slater's condition is not met, every critical point of L identifies a solution to the problem, since KT conditions are still sufficient
- (note: for sufficiency concavity replaced the constraint qualification)
- However, they are no longer necessary
- In the absence of a critical point of L we cannot conclude that a solution to the problem does not exist

Previous example: $f(x) = x, h(x) = -x^2$ Slater's condition fails, lagrangean has no critical points but the solution to the problem exists (x=0)
A failure of Slater's condition is a big deal, but it is somewhat rare

Final Remark

- Note that when all $h_i(x)$ functions are concave, the constraint set D is *convex*, so if the objective function is strictly concave, there is at most one solution to the problem (T7.14).
- Under these circumstances, if you find one critical point of L, you are done!

Exercises: S7.8 (page 198)