Math Camp: Optimization Cheat-Sheet

Jacob Hukill

August 18, 2023

Inequality-Constrained Optimization

Theorem 1 (Kuhn-Tucker Necessary Conditions). *Suppose $x^0 = (x_1^0, ..., x_n^0)$ solves the problem*[1]*:*

$$max \;\; f(x) \;\; subject \; to \;\; g_j(x) \leq c_j, \; j = 1, ..., m$$

where f and $g_1, ..., g_m$ are continuously differentiable functions. Suppose further that the functions g_j corresponding to constraints that are active (binding) at x^0 have linearly independent gradients at x^0 $(NCDQ)$. Then there exist unique numbers $\lambda_1, ..., \lambda_m$ such that:

1. $\frac{\partial f(x_0)}{\partial x_i} - \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x_0)}{\partial x_i} = 0, \quad (i = 1, ..., n)$
2. $\lambda_j \geq 0 \quad (i = 1, ..., n)$
3. $\lambda_j[g_j(x_0) - c] = 0 \quad (i = 1, ..., n)$

Theorem 2 (Sufficient Conditions For Local Maxima). *Let f, $g_1, ..., g_m, h_1, ..., h_k$ be C^2 functions on $\mathbb{R}^n$. Consider the problem of maximizing f on the constraint set:*

$$max \;\; f(x) \;\; subject \; to \;\; g_j(x) \leq b_j, \; j = 1, ..., m; \quad h_i(x) = c_i \; i = 1, ..., k$$

Form the Lagrangian

$$L(x_1, ..., x_n, \lambda_1, ..., \lambda_m, \mu_1, ..., \mu_k) = f(x) - \lambda_1(g_1(x) - b_1) - ... - \mu_k(h_k(x) - c_k)$$

1. Suppose there exist $\lambda_1^*, ..., \lambda_m^*, \mu_1^*, ..., \mu_k^*$ such that the first order conditions are satisfied. That is,

- $\frac{\partial L}{\partial x_1} = 0, ..., \frac{\partial L}{\partial x_n} = 0, \quad at\, (x^*, \lambda^*, \mu^*),$

[1] This could be extended to include equality constraints as well. In which case you would need to include all binding constraints (from equality and inequality constraints) in your constraint qualifications.

- $\lambda_1^* \geq 0, ..., \lambda_m^* \geq 0,$
- $\lambda_1^*(g_1(x^*) - b_1) = 0, ..., \lambda_m^*(g_m(x^*) - b_m) = 0,$
- $h_1(x^*) = c_1, ..., h_k(x^*) = c_k.$

2. For notation's sake, suppose that $g_1, ..., g_e$ are binding at x^* and $g_{e+1}, ..., g_m$ are not binding. Write $(g_1, ..., g_e)$ as g_E. Suppose that the Hessian of L with respect to x at (x^*, λ^*, μ^*) is negative definite on the constraint set

$$\{v : Dg_E(x^*)v = 0 \quad \textit{and} \quad Dh(x^*)v = 0\}$$

Then x^* is a strict local constrained max of f on $C_{g,h}$.

To check this second condition, form the bordered Hessian[2]:

$$\begin{bmatrix} 0 & \cdots & 0 & 0 & \cdots & 0 & \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \frac{\partial g_e}{\partial x_1} & \cdots & \frac{\partial g_e}{\partial x_n} \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \frac{\partial h_k}{\partial x_1} & \cdots & \frac{\partial h_k}{\partial x_n} \\ \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_e}{\partial x_1} & \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_k}{\partial x_1} & \frac{\partial^2 L}{\partial x_1^2} & \cdots & \frac{\partial^2 L}{\partial x_n \partial x_1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_1}{\partial x_n} & \cdots & \frac{\partial g_e}{\partial x_n} & \frac{\partial h_1}{\partial x_n} & \cdots & \frac{\partial h_k}{\partial x_n} & \frac{\partial^2 L}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 L}{\partial x_n^2} \end{bmatrix}$$

- If the last n-(e+k) leading principal minors alternate in sign, with the sign of the determinant of the largest matrix the same as the sign of (-1)^n, then the condition holds.
  - This may seem unwieldy, but consider the case of two x variables and one constraint. So $n-(e+k) = 2-(0+1) = 1$. Therefore we only have to check one principal minor. Furthermore, this principal minor has to be the same sign of (-1)$^n = (-1)^2 > 0$.
- To state the minimization version instead,
  - First transform the problem into a standard form for minimization (i.e. Maximize $-f, etc.$)
  - Change the word maximizing to minimizing

[2]Notice that this is effectively the Hessian of the Lagrangian itself, since constraints that don't bind fall out of the Lagrangian. The difference being that the constraints would be multiplied by (-1); however, this won't change the sign of our determinants since the effect is always being applied so that it cancels itself out.

- Write the inequality constraints as $g_i(x) \geq b_i$,
- Change "negative definite" to "positive definite"
- Change "max" to "min" in the concluding sentence.
- To check the Bordered Hessian test, now we need the largest $n-(e+k)$ principle minors to be the same sign as (-1)$^{e+k}$.

Theorem 3 (Kuhn-Tucker Sufficient Conditions (global)). *Consider the nonlinear programming problem*

$$max \;\; f(x) \;\; subject \ to \;\; g_j(x) \leq c_j, \;\; j=1,...,m$$

where f and $g_1,...,g_m$ are continuously differentiable with f ***concave*** *and $g_1,...,g_m$ all* ***convex****. Suppose that there exist numbers $\lambda_1,...,\lambda_m$ and a feasible vector x_0 such that*

1. $\frac{\partial f(x_0)}{\partial x_i} - \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x_0)}{\partial x_i} = 0, \quad (i=1,...,n)$
2. $\lambda_j \geq 0 \quad (j=1,...,m)$
3. $\lambda_j[g_j(x_0)-c]=0 \quad (j=1,...,m)$

Then x_0 solves the problem.

Theorem 4 (Sufficient Conditions for Quasi-Concave Programming (global)). *Consider the nonlinear programming problem*

$$max \;\; f(x) \;\; subject \ to \;\; g_j(x) \leq c_j, \;\; j=1,...,m$$

where f and $g_1,...,g_m$ are continuously differentiable. Suppose that there exist numbers $\lambda_1,...,\lambda_m$ and a feasible vector x_0 such that

1. $\frac{\partial f(x_0)}{\partial x_i} - \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x_0)}{\partial x_i} = 0, \quad (i=1,...,n)$
2. $\lambda_j \geq 0 \quad (j=1,...,m)$
3. $\lambda_j[g_j(x_0)-c]=0 \quad (j=1,...,m)$
4. $(f_1'(x_0),...,f_n'(x_0)) \neq (0,...0)$
5. *f(x) is quasi-concave, and $\lambda_j g_j(x)$ is quasi-convex for j=1,...,m.*

Then x_0 solves the problem[3].

[3] Condition (4) is a fairly significant addition to the Kuhn-Tucker Sufficient Conditions. It excludes points at which the vector x_0 in f is stationary.

Quasi-concavity in Optimization:

Theorem 5 (Concavity to Quasi-Concavity). *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$. If f is concave on D, then it is also quasi-concave on D. If f is convex on D, then it is also quasi-convex on D.*

Theorem 6 (Quasi-Concavity: Uniqueness, Global Conditions). *Suppose that $f : D \to \mathbb{R}$ be quasi-concave where $D \subset \mathbb{R}^n$ is convex. Then any **local maximum** of f on D is also a **global maximum** of f on D. Moreover, if f is strictly quasi-concave, any local maximum of f is the unique global maximum.*

Theorem 7 (Monotonic Transformations of Quasi-concave functions). *If $f : D \to \mathbb{R}$ is quasi-concave on D, and $F : \mathbb{R} \to \mathbb{R}$ is a monotone non-decreasing function, then the composition $F(f)$ is a quasi-concave function from D to $\mathbb{R}$. In particular, any monotone transform of a concave function results in a quasi-concave function.*

Theorem 8 (Extreme Points Preserved Under Monotonic Transformations). *Suppose a function f has a local maximum (minimum) point given by x_0. If $F : \mathbb{R} \to \mathbb{R}$ is a monotone non-decreasing function, then x_0 will also be a local maximum (minimum) point of the composition $F(f)$. Moreover, global maximums (minimums) of f will be global maximums (minimums) of F(f).*

Common Monotonic Transformations

The following are all common monotonic transformations of the set of positive scalars ++.

- $az + b$, a,b>0
- z^2
- e^z
- $ln(z)$
- z^3

Existence

Theorem 9 (The Weierstrass (Extreme Value) Theorem). *If K is a compact set of $\mathbb{R}^n$, and $f : K \to \mathbb{R}$ is a continuous function, then f($\boldsymbol{x}$) achieves a global maximum on K.*

Approaching an Optimization Problem:

Suppose you have an optimization problem of the form:

$$\max \ \ f(x) \ \ \textit{subject to} \ \ g_j(x) \le b_j, \ j = 1, ..., m; \ \ h_i(x) = c_i \ i = 1, ..., k$$

Steps To Solve:

1. First, we should try to ensure the maximum actually exists. Apply the Weierstrass theorem. If it succeeds, then we are guaranteed to have a solution. If it fails, a solution may not actually exist. However, there still might be an optimum, and we may still be able to characterize it - for example, concavity could mean that a local maximum is a global maximum, and we could still show this using the (bordered) Hessian.

2. Can we transform the problem into something easier to solve? Here, we could possibly transform the problem in a couple ways.

   - We could possibly "reduce" the constrained problem into an unconstrained problem by substituting in the constraint into the objective, if possible. This may make the problem easier to solve. Just make sure that if you solve for the constraint, that you are not implicitly creating assumptions over your variables.
   - Because optima are preserved over an increasing transformation of f, if we can find an "easier" problem to solve by finding such a function g, such that f is a strictly increasing transformation of g, then we can solve the optimization problem using the easier function. This typically involves finding a g that is a concave function, of which has nice properties for global maximization. Just make sure that you are aware of whether the domain is the same across the functions.

3. Now we locate all possible candidates for the solution. We must include all the points that satisfy our FONCs, as well as points that don't meet the requirements to apply the theorem - this means we must also include any points that fail to meet our NCDQ's or possibly non-differentiable at a point[4]. If we do this properly, then if a maximum exists, it must be one of these candidates.

4. Now we can narrow these candidates depending on the structure of the problem. For example, if we have a nonlinear programming problem, then with a concave objective function f and convex constraint functions g, we know that any critical point of the Lagrangian will be a global maximum. First check to see if any of these global conditions (concavity, quasi-concavity) are satisfied. It may be that candidates are automatically global extrema.

[4]Note that in unconstrained optimization, this typically means that we need to include our boundary points, since the FONCs usually apply to interior points

5. If these global conditions are not met, then we need to look closer at each of our candidates. If we have function forms, this involves simply evaluating the objective function at each of our candidates and picking the one that gives the highest value for maximization. Otherwise, we can narrow down potential candidates by describing sufficient conditions (usually over some parameters) such that the points are local maxima. Sometimes we can conclude from this that the local maxima must also then be global maxima (such as if you know an optimum exists and there is only one candidate that is a local maxima).