5 Lecture 5

5.1 Convex Programming[1]

In general, the second order conditions above are a pain to test. It would be nice to know circumstances where we do not have to check them. In the case of unconstrained optimization, we knew that there were various conditions regarding convexity that helped us out - for example, if a function is concave, we know that any point satisfying the first order conditions is a global minimum. It turns out that we don't strictly speaking need full concavity for this result. In fact, we need a weaker property called pseudo-concavity

Definition 8 *A function f is pseudoconcave if*

$$f(x) > f(x^*) \Rightarrow \nabla f\left(x^*\right)'\left(x - x^*\right) > 0$$

Intuitively, this concept is very closely related to that of quasiconcavity (which states that $f(x) \geq f(x^*) \Rightarrow \nabla f\left(x^*\right)'\left(x - x^*\right) \geq 0$) as the following theorem states

Theorem 16 *Let U be a convex subset or $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ be $\mathcal{C}^1$. Then*

1. *If f is pseudoconcave, it is quasiconcave*
2. *If U is open and $\nabla f\left(x\right) \neq 0$, $x \in U$, then f is pseudoconcave if and only if it is quasiconcave*

So, broadly speaking, pseudoconcavity is like quasiconcavity, but rules out 'flat sections'

Also intuitively, pseudoconcavity is enough to make necessary first order conditions also sufficient, first for unconstrained optimization

Theorem 17 *Let U be a convex subset or $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ be $\mathcal{C}^1$ and pseudoconcave. If $\nabla f\left(x^*\right) = 0$ then x^* is a global max of f on U*

[1] We do not have enough time to cover this materiel thoroughly. For more details, see Simon and Blume Chapter 21.4

What about for constrained problems? Here, it is clearly not enough for the objective function to be pseudoconcave, or even concave - it is easy enough to construct a constraint set and a p-concave function such that we obtain a local minimum from the first order conditions (think of the example in the section on equality constraints above) so what to do?

One way out of this is to make the constraint set convex. Intuitively, first order conditions should be sufficient when maximizing a pseudo concave function on a convex set (draw a couple of pictures to convince yourself that this is true). So how do we do this? Well, in the case of inequality constraints, this is easy - we just have to demand that each function is q-convex, so the lower contour sets of each function are convex. The union of convex sets is convex, so we are done. What about equality constraints? Here, the only way to ensure convexity is for the function to be affine (to see this, note that any equality constraint $h(x) = 0$ can be written as two inequality constraints, $h(x) \leq 0$ and $-h(x) \leq 0$. For both of these functions to be q-concave, it must be the case that $h(x)$ is affine).

Putting all this together we get the following theorem:

Theorem 18 *Let U be an open convex subset of $\mathbb{R}^n$, and $f : U \to \mathbb{R}$ be C^1 and pseudoconcave. Let $h_1, ..., h_n$ be a set of affine functions mapping U to $\mathbb{R}$, and $g_1...g_p$ be a set of C^1, quasi-convex functions mapping U to $\mathbb{R}$. Consider the problem*

- *Find $x \in \mathbb{R}^n$*
- *In order to maximize $f(x)$*
- *Subject to*

$$\begin{aligned} h(x) &= 0 \\ g(x) &\leq 0 \end{aligned}$$

Suppose <SOME CONSTRAINT QUALIFICATIONS HOLD>, and for some $x^* \in U$ there exists $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p$ that satisfy the KKT conditions, then x^* is the global maximum of the constrained problem.

So these assumptions have bought us a lot. Not only are the first order conditions now sufficient, but the also identify global, not local maxima. Hurrah!

To complete the theorem, we need to identify some reasonable constraint qualifications. We could use the regularity conditions that we have assumed up to now, but in convex programming problems there is a nicer set of qualifications which will do:

Definition 9 *A set of constraints satisfy the Slater Constraint qualification if g_i is convex all i, and there exists an $\bar{x}$ such that $h(\bar{x}) = 0$ and $g(\bar{x}) < 0$*

5.2 Sensitivity Analysis

One nice feature of the Kuhn Tucker approach is that the leGrange multipliers λ and μ have a nice, intuitive interpretation: they tell us how the value of the constrained optimization problem changes are we change the constraints. In order to make this statement formal, we will first of all think about equality constraints, and define a paramerterized version of the constrained problem:

Definition 10 *Let $\mathcal{P}(c)$ refer to the parameterized problem of maximizing a function $f(x)$ relative to a set of constraints*

$$h(x) = c$$

where $c \in \mathbb{R}^m$ is a vector of parameters.

Our claim is that the Lagrangian multipliers at a local solution to this problem tell us something about how the value of that local solution changes as we change the parameter vector c (we will concentrate on perturbations around $c = 0$.

Theorem 19 *Suppose that, when $c = 0$, x^* is a local maximizer and a regular point of $x^* = 0$. Let λ^* be a set of LaGrange multipliers such that (x^*, λ^*) satisfy the KKT conditions. Assume that $\Psi(x^*, \lambda)$ is negative definite in $T = \ker(h'(x^*))$. Then there exists an $\varepsilon > 0$, $B(o, \varepsilon) \in \mathbb{R}^m$ and $x : B(o, \varepsilon) \to \mathbb{R}^n$ continuous such that $x(c)$ is the strict local maximum for $\mathcal{P}(c)$, and*

$$\frac{d}{dc_i} \left(f(x(c)) \right|_{c=0} = \lambda_i^*$$

Proof (Sketch). *Consider the following system of equations*

$$\begin{bmatrix} \nabla f(x) + \sum \lambda_i h_i(x) \\ h(x) - c \end{bmatrix} = \Phi(x, \lambda, c) = 0$$

where $\Phi : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^{n+m}$. *Note that* $\Phi(x^*, \lambda^*, 0) = 0$. *A claim, that we will not prove, is that regularity is enough to ensure that* $\Phi_{x,\lambda}(x^*, \lambda^*, 0)$ *is of full rank. Therefore, by the implicit function theorem, there exists a continuous function*

$$\begin{pmatrix} x \\ \lambda \end{pmatrix} : B(o, \varepsilon) \to \mathbb{R}^{n+m}$$

such that $x(0) = x^*,\ \lambda(0) = \lambda^*$ *and such that* $\Phi(x(c), \lambda(c), c) = 0$. *This means that* $x(c), \lambda(c)$ *satisfy the first order conditions for the problem* $\mathcal{P}(c)$. *We can also show that* $\Psi(x(c), \lambda(c))$ *is negative semi-definite, so* $x(c)$ *is a local maximizer of* $\mathcal{P}(c)$.

Now, we know that

$$\begin{aligned}
\sum_i \frac{\partial f(x(c))}{\partial x_i} \frac{\partial x_i(c)}{\partial c_n} + \sum_m \left(\lambda_m^* \sum_i \frac{\partial h_m(x(c))}{\partial x_i(c)} \frac{\partial x_i(c)}{\partial c_n} \right) &= 0 \\
\sum_i \frac{\partial h_m(x(c))}{\partial x_i(c)} \frac{\partial x_i(c)}{\partial c_n} &= 0 \ m \neq n \\
\sum_i \frac{\partial h_n(x(c))}{\partial x_i(c)} \frac{\partial x_i(c)}{\partial c_n} &= 1 \\
\frac{\partial f(x(c))}{\partial c_n} &= \sum_i \frac{\partial f(x(c))}{\partial x_i} \frac{\partial x_i(c)}{\partial c_n} = \lambda_n^*
\end{aligned}$$

■

- example (box)
- example (farmer)