Advanced Mathematical Programming IE417

Lecture 7

Dr. Ted Ralphs

Reading for This Lecture

- Chapter 4, Section 2

Optimality Conditions

Inequality Constrained Problems (continued)

Fritz-John Necessary Conditions

Theorem 1. *Consider the feasible region* $S = \{x \in X : g_i(x) \leq 0, i \in [1, m]\}$ *where* X *is a nonempty open set in* $\mathbb{R}^n$ *and* $g_i : \mathbb{R}^n \to \mathbb{R}, i \in [1, m]$. *Given a feasible* $x^* \in S$, *set* $I = \{i : g_i(x^*) = 0\}$. *Assume that* f *and* g_i *are differentiable at* x^* *for* $i \in I$ *and* g_i *is continuous at* x^* *for* $i \notin I$. *If* x^* *is a local minimum, then there exists* $\mu \in \mathbb{R}^m$ *such that*

$$\begin{aligned} \mu_0 \bigtriangledown f(x^*) + \sum \mu_i \bigtriangledown g_i(x^*) &= 0 \\ \mu_i g_i(x^*) &= 0 \, \forall i \in [1, m] \\ \mu &\geq 0 \\ \mu &\neq 0 \end{aligned}$$

Terminology

- The μ_i's are called *Lagrange multipliers* or *dual multipliers*.
- The requirement that $x^* \in S$ is called the *primal feasibility* (PF) condition.
- The requirement that $\mu_0 \bigtriangledown f(x^*) + \sum \mu_i \bigtriangledown (x^*) = 0$ is called the *dual feasibility* (DF) condition.
- The requirement that $\mu_i g_i(x^*) = 0 \ \forall i \in [1, m]$ is called the *complementary slackness* (CS) condition.
- *FJ points* are those satisfying PF, DF and CS.

Fritz-John Sufficient Conditions

- Note that a point is an FJ point if and only if $F_0 \cap G_0$ is empty.
- Notation and setup as for necessary conditions.

Theorem 2. *If there exists $N_\epsilon(x^*), \epsilon > 0$ such that f is pseudoconvex and $g_i, i \in I$ are strictly pseudoconvex over $N_\epsilon(x^*) \cap S$, where S is the relaxed feasible region without the nonbinding constraints, then x^* is a local minimum.*

- There are also other possible sufficient conditions.

Remarks on the FJ conditions

- These conditions hold trivially in many cases.
- In particular, if $G_0 = \emptyset$, they will hold, regardless of the objective function (take $\mu_0 = 0$).
- Even for LP, there are non-optimal FJ points.
- We want to force $\mu_0 > 0$ in order to take the objective function into account.
- We do this by using a *constraint qualification*.

KKT Necessary Conditions

- We now require that the gradients of the binding constraints be linearly independent. This implies that $G_0 \neq \emptyset$ and hence $\mu_0 > 0$.
- In this case, we can drop the condition that $\mu \neq 0$ and we get x^* locally optimal $\Rightarrow$ there exists $\mu \in \mathbb{R}^m$ such that

$$\begin{aligned} \nabla f(x^*) + \sum \mu_i \nabla g_i(x^*) &= 0 \\ \mu_i g_i(x^*) &= 0 \, \forall i \in [1, m] \\ \mu &\geq 0 \end{aligned}$$

Remarks on the KKT conditions

- Again, we have PF, DF and CS conditions. These make up the *KKT conditions*.
- x^* is a *KKT point* if the KKT conditions are satisfied at x^*.
- For a linear program, the KKT conditions are simply the standard optimality conditions for LP.
- Using previous notation, note that x^* is a KKT point if and only $F_0 \cap G_0'$ is empty
- Furthermore, x^* is a KKT point if and only if x^* is the solution to the first-order LP approximation to the NLP

$$\min\{f(x^*) + \bigtriangledown f(x^*)^T(x - x^*) : g_i(x^*) + \bigtriangledown g_i(x^*)^T(x - x^*) \leq 0, i \in [1, m]\}$$

KKT Sufficient Conditions (1^{st} Order)

- We have the same setup as before.

  Theorem 3. *Let* x^* *be a KKT point and let* $I = \{i : g_i(x^*) = 0\}$*. If* f *is pseudoconvex at* x^* *and if* $g_i, i \in I$ *are differentiable and quasiconvex at* x^**,* x^* *is then a global optimal solution.*

- There are other possible sufficient conditions.