15.084J Recitation Handout 4

Fourth Week in a Nutshell:

- Separating Hyperplanes
- Theorem of the Alternative (Farkas Lemma)
- Necessary Conditions for Optimum of Constrained Problem
- Finding Optima

Separating Hyperplanes

Main point: Two closed, convex, disjoint sets can be separated by a hyperplane.

Fiddly details: OK, only the interiors need to be disjoint. And they only need to be relatively convex. And it works for the closure of open sets. And sometimes we can do varying levels of separation....

A hyperplane H separates S_1 and S_2 if $S_1 \subseteq H^+$ and $S_2 \subseteq H^-$. It "properly separates" them if their intersection does not touch the hyperplane. It "strictly separates" them if neither one touches the hyperplane. It "strongly separates" them if you could move the hyperplane by $\pm\epsilon$ perpendicular to itself and still separate them.

Why do we care? A closed convex set S and a point not in S can be strongly separated. Any closed convex set can be defined as an intersection of halfspaces. The CH of a non-convex set is the intersection of all halfspaces containing it. Two closed convex sets can be strongly separated.

Farkas Lemma/Theorem of the Alternative Remember that great lemma about LP duality, where either the primal was feasible, or the dual was feasible? We can apply a similar argument to show that either an improving direction exists, or we've got an equation involving gradients that holds. This leads us to the necessary conditions.

Farkas' Lemma: Either $Ax \leq 0, c^t x > 0$ has a solution, or $A^t y = c, y \geq 0$ does. We can slightly generalize this to: Either $Ax < 0, Bx \leq 0, Hx = 0$ or $A^t u + B^t v + H^t w, u \geq 0, v \geq 0, e^t u = 1$ has a solution. The first gives us a direction of improvement, or the second demonstrates that all first-order improvements are blocked.

Necessary Conditions For Optimality

Just like in unconstrained optimization, there's really only one criterion: there is no locally improving direction. Just like in unconstrained, this is *necessary*, but not sufficient.

How can we not have a locally improving direction? Either we're in the interior, and the unconstrained condition of gradient being zero holds, or we have tight constraints which block all movement in directions the gradient says are good.

And how can they do that? Well, for a single tight inequality constraint, if it's gradient points directly opposite the gradient of the function, it's blocking all locally improving directions. For multiple tight constraints, if a positive linear combination can directly oppose the gradient of the function, that works too.

And equality constraints? They're the same as a pair of inequality constraints.

In equation form, $\nabla f(x) + \sum_i \lambda_i \nabla g_i(x) = 0$, with $\lambda_i \geq 0$. An equality constraint appears twice with opposite signs, or alternatly appears once with its λ being unconstrained.

Finding Optima

OK, how to find optimum value in practice using KKT conditions?

For every subset (including the empty set) of inequality constraints, use the above equation including all equality constraints and the chosen inequality constraints to find candidate points. Check whether they are feasible for the other constraints. If you've got more than one point left, check $f(x_1)$ and $f(x_2)$ to see which one has a lower value (you can rule out some local minima this way). If your feasible region is unbounded, check what happens as x runs off to infinity.

When does this fail? When the constraints are doing something special – your feasible region is collapsing to a point or a ray. Or when your constraints are non-differentiable (hopefully this happens at a limited number of points, and you can check them directly). Or when you have an infinite collection of constraints.

And what if you can't find gradients in closed form, or have lots of constraints? Tune in next week, when we cover actual algorithms...