Advanced Mathematical Programming
IE417

Lecture 20

Dr. Ted Ralphs

Reading for this lecture

- Sections 9.3

Exact Penalty Functions

- The penalty functions

$$\begin{aligned}\phi(g_i(x)) &= \max\{0, g_i(x)\}, \text{and} \\ \psi(h_i(x)) &= |h_i(x)|\end{aligned}$$

  are called *exact penalty functions*.

- Exact penalty functions achieve optimality for a finite value of μ.
- For a convex program, μ must exceed the maximum of the absolute values of the optimal Lagrange multipliers.

Augmented Lagrangian Methods

- The problem with the foregoing penalty functions is that they are not differentiable everywhere.
- Idea: Shift the penalty term a little in order to reduce its dominance of the objective function.
- Consider the penalty function $\psi(h_i(x)) = [h_i(x) - \Theta_i]^2$.
- Assuming only equality constraints, the penalized objective function can then be written as

$$F(x,v) = f(x) + \Sigma v_i h_i(x) + \mu \Sigma [h_i(x)]^2$$

Augmented Lagrangian Methods (continued)

- Note that if (x^*, v^*) is a KKT point for the original problem, then in fact

$$\nabla F(x^*, v^*) = \nabla f(x^*) + \Sigma v_i^* \nabla h_i(x^*) + 2\mu\Sigma[h_i(x^*)]\nabla h_i(x^*) = 0$$

- All we then need is for (x^*, v^*) to be a local minimum.
- The quadratic penalty serves this purpose by convexifying the objective function.

Another Interpretation

- We can view the augmented Lagrangian (ALAG) in two other ways.
- It is the usual quadratic penalty function with respect to

$$\min\{f(x) + \Sigma v_i h_i(x) : h_i(x) = 0 \ \forall i\}$$

- It is also the Lagrangian for

$$\min\{f(x) + \mu\Sigma[h_i(x)]^2 : h_i(x) = 0 \ \forall i\}$$

Magnitude of the Penalty

- If (x^*, v^*) is a KKT solution satisfying the second-order sufficiency conditions, then there exists a μ^* such that for $\mu \geq \mu^*$, then x^* is a strict local minimum for $F(\cdot, v^*)$.
- Intuitively, needs to be big enough to "convexify" the function $F(x, v)$.
- Hence, if f is already convex, any non-negative multiplier will do.
- Note that in practice, we will use a separate penalty multiplier for each constraint.

Implementation of ALAG

- Select some initial Lagrange multipliers v^* and define $VIOL(x) = \max\{|h_i(x)| \forall i\}$. Set $VIOL(x_0) = \infty$.
- Inner Loop
  - Minimize $F(x, v^*)$ to obtain x_k. If $VIOL(x_k) < \varepsilon$, STOP.
  - If $VIOL(x_k) \leq (0.25)VIOL(x_{k-1})$, exit the inner loop.
  - Otherwise, multiply the penalty on each constraint for which $|h_i(x_k)| > (0.25)VIOL(x_{k-1})$ by 10.
- Outer Loop Update: Set $v_i^* \leftarrow v_i^* + 2\mu_i h(x_k)$.

Comments on ALAG Methods

- Note that the similarity to solving the Lagrangian dual, except that we now get both primal and dual solutions through the use of the penalty term.
- Because of the Lagrangian augmentation, we are able converge without a large penalty.
- This eliminates the ill-conditioning effects of the standard penalty method.
- This procedure is very much akin to solving the Lagrangian dual of the quadratic penalty function using steepest ascent with a fixed step length.
- We could deflect the ascent direction using something akin to a quasi-Newton method if the penalty multipliers do get large.

Extending ALAG Methods

- Up until now, we have assumed equality constriants.
- We can extend to include the case where there are also inequalities by replacing $g_i(x) \leq 0$ with $g_i(x) + s_i^2 = 0$.
- Plugging this into the function $F(x, v)$ from before, we obtain

$$
\begin{aligned}
F(x, u, v) \quad = \quad & f(x) + \mu \sum_{i=1}^{m} \max^2\{g_i(x) + u_i/2\mu, 0\} \\
& - \sum_{i=1}^{m} u_i^2/4\mu + \sum_{i=1}^{m} v_i h_i(x) + \mu \sum_{i=1}^{m} h_i(x)^2
\end{aligned}
$$