CH2.3 AUGMENTED LAGRANGIANS

Monday, April 1, 2024 7:22 PM

- Augmented method adds robustness to the dual ascent method.
- Yields convergence without assuming strict convexity or finiteness of objective function f

LAGRANGIAN: $L(x,y) = f(x) + y^T(Ax-b)$

AUGMENTED LAGRANGIAN: $L_\rho(x,y) = f(x) + y^T(Ax-b) + \left(\frac{\rho}{2}\right)\|Ax-b\|_2^2$

$\rho > 0$: "penalty parameter"

If $\rho = 0$, then Augmented is equal to original Lagrangian.

Augmented Problem: minimize: $f(x) + \left(\frac{\rho}{2}\right)\|Ax-b\|_2^2$
subject to: $Ax = b$

Augmented Dual: $g_\rho(y) = \inf_x L_\rho(x,y)$

Augmented Algorithm: "METHOD OF MULTIPLIERS" for solving (2.1)

$$\begin{cases} x^{k+1} := \operatorname{argmin}_x L_\rho(x, y^k) \rightarrow \text{x-minimization step} \\ y^{k+1} := y^k + \rho(Ax^{k+1} - b) \rightarrow \text{multiplier update step} \end{cases}$$

$$(2.1)\begin{cases} \text{minimize } f(x) \\ \text{subject to } Ax = b \end{cases}$$

This is the same as the standard Dual ascent, except that the x-minimization step uses the Augmented Lagrangian, and uses the penalty parameter p instead of step size alpha^k.
The Method of Multipliers (MoM) converges under more general conditions than dual ascent, even in cases when f takes on +infinity or **is not strictly convex**.
MoM does **not even require f to be differentiable**.

Choice of penalty parameter ρ:

Assume f differentiable.

$Ax^* - b = 0, \quad \nabla f(x^*) + A^T y^* = 0$, (primal and dual feasibility)

x^{k+1} minimizes $L_\rho(x, y^k)$

$\hookrightarrow 0 = \nabla_x L_\rho(x^{k+1}, y^k)$

k^{th}-iteration of y, transposed

$$0 = \nabla_x\left(f(x^{k+1}) + y^{kT}(Ax^{k+1} - b) + \left(\frac{\rho}{2}\right)\|Ax^{k+1} - b\|_2^2\right)$$

k+1 iteration, NOT power!

$$0 = \nabla_x f(x^{k+1}) + y^{kT}\nabla_x Ax^{k+1} + \nabla_x \frac{\rho}{2}(Ax^{k+1} - b)^T(Ax^{k+1} - b)$$

$$= \frac{\rho}{2}\nabla_x\left(x^TA^TAx - x^TA^Tb - b^TAx + b^Tb\right)$$

$$= \frac{\rho}{2}\nabla_x\left(x^TA^TAx - 2x^TA^Tb + b^Tb\right)$$

$$= \frac{\rho}{2}\left(2A^TAx - 2A^Tb\right)$$

$$0 = \nabla_x f(x^{k+1}) + A^Ty^k + \rho(A^TAx - 2A^Tb)$$

$$0 = \nabla_x f(x^{k+1}) + A^Ty^k + \rho A^T(Ax - 2b)$$

$$0 = \nabla_x f(x^{k+1}) + A^T(\underbrace{y^k + \rho(Ax - 2b)}_{= y^{k+1}})$$

$$0 = \nabla_x f(x^{k+1}) + A^T y^{k+1}$$