Using a dual problem to solve a constrained primal problem

To solve: $\min f(x)$ st $Ax=b$

Lagrangian $\mathcal{L}(x,y) = f(x) + \langle y, Ax-b \rangle$

Dual function $g(y) = \inf_x \mathcal{L}(x,y)$

Dual problem $\sup_y g(y)$

If y^* is dual optimal, find primal optimal x^* by

$$x^* = \operatorname{argmin}_x \mathcal{L}(x, y^*)$$

Dual Ascent method

Idea: run gradient ascent on dual problem

Need: $\nabla_y g(y)$

$$g(y) = \inf_x f(x) + \langle y, Ax - b \rangle$$

$$\nabla_y g(y) = \nabla_y \inf_x f(x) + \langle y, Ax - b \rangle$$

$$= Ax^* - b \quad \text{where } x^* \text{ is minimizer} \quad x^* = \arg\min_x \mathcal{L}(x, y)$$

Dual ascent method:

$$x^{k+1} = \arg\min_x \mathcal{L}(x, y^k) \quad \longleftarrow x \text{ minimization}$$

$$y^{k+1} = y^k + \alpha^k (Ax^{k+1} - b) \quad \longleftarrow \text{dual ascent}$$

If $f(x)$ separates, computation distributes

eg $f(x) = \sum_i |x_i|$

$$\arg\min_x \mathcal{L}(x, y^k) = \arg\min \sum_i \left(|x_i| + \langle y, a_i \rangle \right)$$ w/ a_i a col of

Method of Multipliers

To solve: $\min f(x)$ st $Ax=b$

Augmented Lagrangian $(\rho > 0)$

$$\mathcal{L}_\rho(x,y) = f(x) + \langle y, Ax-b \rangle + \frac{\rho}{2} \|Ax-b\|_2^2$$

Method of multipliers:

$$x^{k+1} = \min_x \mathcal{L}_\rho(x, y^k)$$

$$y^{k+1} = y^k + \rho(Ax^{k+1} - b)$$

note particular step size

Optimality Conditions

$$\nabla_x \mathcal{L} = 0 \Rightarrow \nabla f(x^*) + A^t y^* = 0 \quad \text{— dual feasibility}$$

$$\nabla_y \mathcal{L} = 0 \Rightarrow Ax^* - b = 0 \quad \text{— primal feasibility}$$

Claim: Each (x^{k+1}, y^{k+1}) is dual feasible.

As x^{k+1} minimizes $f(x) + \langle y^k, Ax-b \rangle + \frac{\rho}{2} \| Ax-b \|_2^2$

$$\Rightarrow 0 = \nabla f(x^{k+1}) + A^t y^k + \rho A^t (Ax^{k+1} - b)$$

$$= \nabla f(x^{k+1}) + A^t (y^k + \rho(Ax^{k+1} - b))$$

$$= \nabla f(x^{k+1}) + A^t y^{k+1} \Rightarrow \text{dual feasibility}$$

Claim: Primal feasibility achieved as $k \to \infty$.

Alternating direction method of multipliers (ADMM)

$$\min \quad f(x) + g(z) \quad \text{st} \quad Ax + Bz = c$$

Augmented lagrangian

$$\mathcal{L}_\rho(x, z, y) = f(x) + g(z) + \langle y, Ax + Bz - c \rangle + \frac{\rho}{2} \| Ax + Bz - c \|_2^2$$

ADMM

$$x^{k+1} = \underset{x}{\operatorname{argmin}} \ \mathcal{L}_\rho(x, z^k, y^k) \qquad x \text{ min}$$

$$z^{k+1} = \underset{y}{\operatorname{argmin}} \ \mathcal{L}_\rho(x^{k+1}, z, y^k) \qquad z \text{ min}$$

$$y^{k+1} = y^k + \rho \left(A x^{k+1} + B z^{k+1} - c \right) \qquad \text{dual ascent}$$

~~Criteria:~~

Optimality conditions

$$\nabla_y \mathcal{L} = 0 \Rightarrow A x^* + B z^* - c = 0 \quad \leftarrow \text{satisfied in limit}$$

$$\nabla_x \mathcal{L} = 0 \Rightarrow \nabla f(x^*) + A^t y^* = 0 \quad \leftarrow \text{satisfied in limit}$$

$$\nabla_z \mathcal{L} = 0 \Rightarrow \nabla g(z^*) + B^t \cancel{z} y^* = 0 \quad \leftarrow \text{Solved at each step}$$

Lasso (ℓ_2 - penalized ℓ_1 minimization)

$$\min \quad \lambda \|x\|_1 + \frac{1}{2}\|Ax-b\|_2^2$$

ADMM form

$$\min \quad \frac{1}{2}\|Ax-b\|_2^2 + \lambda\|z\|_1 \qquad \text{st} \quad x - z = 0$$

$$\mathcal{L}_\rho = \frac{1}{2}\|Ax-b\|_2^2 + \lambda\|z\|_1 + \langle y, x-z\rangle + \frac{\rho}{2}\|x-z\|_2^2$$

ADMM:

$$x^{k+1} = \underset{x}{\operatorname{argmin}} \ \frac{1}{2}\|Ax-b\|_2^2 + \langle y^k, x\rangle + \frac{\rho}{2}\|x-z^k\|_2^2 \rightarrow x^{k+1} = (A^tA+\rho I)^{-1}(A^tb+\rho z^k - y^k)$$

$$z^{k+1} = \underset{z}{\operatorname{argmin}} \ \lambda\|z\|_1 - \langle y^k, z\rangle + \frac{\rho}{2}\|x^{k+1}-z\|_2^2 \rightarrow z^{k+1} = \text{Soft Thresh}_{\lambda/\rho}\left(x^{k+1} + y^k/\rho\right)$$

$$y^{k+1} = y^k + \rho(x^{k+1} - z^{k+1})$$

ADMM works under few assumptions (f, g convex but not diff.able)

Distributing on z