Convex Optimization   Alternating Direction Method of Multipliers

Problem 1.

$$\min \ \frac{1}{2}x^T H x + h^T x$$
$$\text{s.t.} \quad Ax = b$$
$$l \le Cx \le u$$

a)

reformulate as $\min f(x) + g(Lx)$

where $f(x) = \frac{1}{2}x^T H x + h^T x + I_{Ax=b}(x)$.

i) $L = C$. then. $g(Lx) = I_{[l,u]}(Cx)$, $g(x) = I_{[l,u]}(x)$

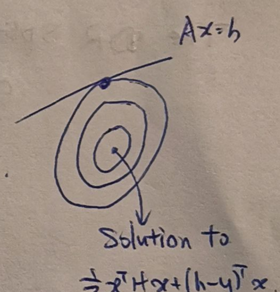
ii) $L = I$. then. $g(Lx) = I_{l \le Cx \le u}(x)$, $g(x) = I_{l \le Cx \le u}(x)$.

Dual problem:

$$\min \ f^*(-L^T\mu) + g^*(\mu).$$

$$f^*(y) = \sup_x y^T x - f(x) = \sup_x y^T x - \left[\frac{1}{2}x^T H x + h^T x + I_{Ax=b}(x)\right]$$

$$= -\inf_x \frac{1}{2}x^T H x + (h-y)^T x + I_{Ax=b}(x).$$



this is actually another QP problem with equation constraints.

Just the same as primal problem.

$$g^*(y) = \sup_x y^T x - g(x)$$

for $L = C$: $g^*(y) = \sup_x y^T x - I_{[l,u]}(x)$

$$= \sup_x y^T x$$
$$\text{s.t.} \quad x \in [l,u]^n.$$

this is a linear Programming Problem, the optimal is chosen at its vertices: for $[l,u]^n$, there are 2^n vertices.

for $L = I$. $g^*(y) = \sup_x y^T x$

$$\text{s.t.} \quad l \le Cx \le u$$

If C is sparse, then the Polytope $\{x \mid l \le Cx \le u\}$ has possibly smaller number of vertices, therefore easier to find the solution.

for $L=C$, $f^*(-L^T\mu) = -\inf_x \frac{1}{2}x^T H x + (h + L^T\mu)^T x$

s.t $Ax=b$

for $L=I$, $f^*(-I\mu) = -\inf_x \frac{1}{2}x^T H x + (h+\mu)^T x$

s.t $Ax=b$

b)

Because both primal and dual problems are convex, with Both $L=I, C$

Hence the primal and dual problems can both be solved by DR splitting.

For primal problem,

$$x^k = \text{prox}_{\gamma f}(z^k)$$
$$y^k = \text{prox}_{\gamma g(L)}(2x^k - z^k)$$
$$z^{k+1} = z^k + 2\alpha(y^k - x^k).$$

While for the dual problem

$$\mu^k = \text{prox}_{\gamma g^*}(v^k)$$
$$\eta^k = \text{prox}_{\gamma(f^*\circ(-L^T))}(2\mu^k - v^k)$$
$$v^{k+1} = v^k + 2\alpha(\eta^k - \mu^k).$$

By Moreau's identity.

$$\text{prox}_{\gamma g^*} = I_d - \text{prox}_{\gamma g}.$$

$$\text{prox}_{\gamma(f^*\circ(-L^T))}(2\mu^k - v^k) = 2\mu^k - v^k + \gamma . L . x^*,$$

$$x^* \in \arg\min_x \{ f(x) + \frac{\gamma}{2} \| -Lx - \gamma^{-1}(2\mu^k - v^k) \|_2^2 \} \quad . \quad (1)$$

first compare D_1 and D_2.

then μ^k update is dependent on g.

$\text{prox}_{\gamma g(I)} \to \Pi_I$. $\text{prox}_{\gamma g(C)} \to \Pi_C$.

Normally, projection onto a polyhedro is difficult,

but if C is very sparse, for example $C = \begin{bmatrix} 1 & 0 \\ 0 & \cdots & 0 \end{bmatrix}$

the projection onto this polyhedron might be easier.

the η^k update. If L is sparse, from (1) we know this computation would be easier. again image $C = \begin{bmatrix} 1 & 0 & \cdots \\ 0 & 0 & \\ 0 & & 0 \end{bmatrix}$.

therefore, betwen D_1, and D_2, I would choose $L = C$.

- Compare primal and dual

for the update regarding ~~x~~ $f(x)$, again the dual with $L = C$ wins.

for the update regarding $g(x)$, we will always use.

Moreau's Identity.

$$\mathrm{prox}_{\gamma g^*} = \mathrm{Id} - \mathrm{prox}_{\gamma g}.$$

And with $\mathrm{prox}_{\gamma g(C)} = \mathrm{prox}_{\gamma \cdot I_{l \le Cx \le u}}$.

$$\mathrm{prox}_{\gamma (g \circ L)} = \mathrm{prox}_{\gamma \cdot I_{l \le Cx \le u}}.$$

this $g(x)$ related update is essentially the same, and are both better when $L = C$ is sparse.

therefore, I would choose <u>dual</u> with $L = C$.

Problem 2.

Consider

$$\max \ -\frac{1}{2}x^T P x - S_C(y)$$
$$\text{s.t.} \quad -Px + A^T y = -q$$
$$y \in (C^\infty)^\circ.$$

$P \in S_+^n$

C closed convex set

If there exists some $\bar{x}$. such that.

$P\bar{x} = 0$. $A\bar{x} \in C^\infty$. and $q^T\bar{x} < 0$

then the dual problem is infeasible.

proof:

$$\inf_y \{ y^T A\bar{x} \mid -Px + A^T y = -q \}$$
$$= \inf_y \{ (-q^T + x^T P)\bar{x} \mid -Px + A^T y = -q \}$$
$$= \inf_y \{ -q^T\bar{x} \mid -Px + A^T y = -q \} > 0$$

$$\sup_y \{ y^T A\bar{x} \mid y \in (C^\infty)^\circ \} \Longleftarrow$$

by definition of polar cone, $y^T A\bar{x} = y^T(A\bar{x}) \le 0$ Because $A\bar{x} \in C^\infty$.

therefore. $\sup_y \{ y^T A\bar{x} \mid y \in (C^\infty)^\circ \} \le 0$.

$\Rightarrow$ $\inf_y \{ y^T A\bar{x} \mid -Px + A^T y = -q \} > \sup_y \{ y^T A\bar{x} \mid y \in (C^\infty)^\circ \}$

therefore. $\{ y \mid -Px + A^T y = -q \} \cap \{ y \mid y \in (C^\infty)^\circ \} = \phi$.

Problem 3

$$\min_{x \in \mathbb{R}^n} \|x\|_1 \qquad A \in \mathbb{R}^{m\times n}.\ \text{full row rank}.$$

$$\text{s.t.} \quad Ax = b \qquad b \in \mathbb{R}^m$$

the equivalent way using indicator function

$$\min_{x \in \mathbb{R}^n} \|x\|_1 + I_C(z).$$

$$\text{s.t.} \quad Ax = z$$

$$z \in \{b\} =: C.$$

Therefore, the ADMM update.

$$x^{k+1} \in \arg\min_x \left(\|x\|_1 + \frac{\rho}{2} \| Ax - z^k + \frac{1}{\rho} y^k \|_2^2 \right)$$

this is a Lasso problem ; and has a closed form. for x^{k+1}

$$\cancel{\rho A^T (Ax - z^k + \frac{1}{\rho} y^k) + \mathrm{sgn}(x)}$$

$$z^{k+1} = \Pi_C \left(Ax^{k+1} - z + \frac{1}{\rho} y^k \right) = b.$$

$$y^{k+1} = y^k + \rho \cdot \left(Ax^{k+1} - z^{k+1} \right).$$