HW3.1.

Let $w \in \underset{z}{\operatorname{argmin}}\, L(z, \lambda)$. Show that $Aw - b \in \partial g(\lambda)$.

$g(\lambda) = -f^*(-A^T\lambda) - \lambda^T b$, where $-f^*(y) = \min_x f(x) - y^T x$, is the conjugate func.

$\partial g(\lambda) = A \partial f^*(-A^T\lambda) - b$. (Due to subgradient property).

Next we show that $w \in \partial f^*(-A^T\lambda)$.

This is equivalent to $-A^T\lambda \in \partial f(w)$

$$\Longleftrightarrow w \in \underset{z}{\operatorname{argmin}}\, f(z) - (-A^T\lambda)^T z$$

$$\boxed{\begin{array}{c} x \in \partial f^*(y) \\ \Updownarrow \\ y \in \partial f(x) \\ \Updownarrow \\ x \in \operatorname{argmin} f(z) - y^T z \end{array}}$$

Since $w \in \underset{z}{\operatorname{argmin}}\, L(z, \lambda) = \underset{z}{\operatorname{argmin}}\, f(z) + \lambda^T(Az) - \lambda^T b$.

Hence. $w \in \underset{z}{\operatorname{argmin}}\, f(z) + \lambda^T A z$.

$\Rightarrow w \in \partial f^*(-A^T\lambda)$.

$\Rightarrow Aw - b \in \partial g(\lambda)$.

Thm 1: f is L-smooth $\Rightarrow f^*$ is $1/L$ strongly convex

Proof: f is Lsmooth $\Rightarrow \nabla f$ is Lipschitz

$$\Rightarrow (\nabla f(x) - \nabla f(y))^T (x-y) \geq \tfrac{1}{L} \|\nabla f(x) - \nabla f(y)\|^2$$

Let $u = \nabla f(x)$, $v = \nabla f(y)$

$$\Rightarrow x \in \partial f^*(u), \quad y \in \nabla f^*(v)$$

$$\Rightarrow (u-v)^T (x-y) \geq \tfrac{1}{L} \|u-v\|^2$$

$$\forall u \in \partial f^*(x), v \in \partial f^*(y)$$

$$\Rightarrow f^* \text{ is } \tfrac{1}{L} \text{ strongly convex}$$

Thm 2: f is M-strongly convex $\Rightarrow f^*$ is $\frac{1}{M}$-smooth

Proof:

Let $u = \nabla f(x_u)$, $v = \nabla f(x_v)$

$$\Rightarrow x_u = \nabla f^*(u) = \underset{z}{\operatorname{argmin}} (f(z) - u^T z)$$
$$x_v = \nabla f^*(v) = \underset{z}{\operatorname{argmin}} (f(z) - v^T z)$$

(Because of closeness & strong convexity)

That is, x_u is the minimizer of the strongly convex function $f(z) - u^T z$

$$\Rightarrow f(x_v) - u^T x_v \geq f(x_u) - u^T x_u + \frac{M}{2} \|x_v - x_u\|^2 \quad \forall x_v$$

Now changing the roles of x_v and x_u, after noting that x_v is the minimizer of $f(z) - v^T z$,

$$f(x_u) - v^T x_u \geq f(x_v) - v^T x_v + M/2 \|x_v - x_u\|^2$$

Adding the two $\Rightarrow$

$$f(x_v) + f(x_u) + v^T x_u - u^T x_v \geq f(x_v) + f(x_u) + M\|x_v - x_u\|^2$$

$$\Rightarrow M\|x_v - x_u\|^2 \leq u^T(x_u - x_v) + v^T(x_v - x_u)$$

$$= (u^T - v^T)(x_u - x_v)$$

$$\leq \|u^T - v^T\| \|x_u - x_v\|$$

$$\Rightarrow \|x_v - x_u\| \leq \frac{1}{M}(u - v)$$

$$\Rightarrow f^* \text{ is } \frac{1}{M}\text{-smooth}$$

Conclusion 1: From thm-1 and thm-2, if f is L-smooth and M-convex, f^* is $\frac{1}{M}$ smooth & $\frac{1}{L}$-strongly convex

Conclusion 2: In gradient dual ascent of f, we do the GD of f^* which is $\frac{1}{M}$-smooth & $\frac{1}{L}$-convex $\Rightarrow$ Standard convergence results apply

$$\Rightarrow \cancel{f(x_k)} \|x_k - x^*\|^2 \leq \left(1 - \frac{2}{1 + \frac{1/M}{1/L}}\right)^{2k} \|x_0 - x^*\|^2$$

$$\Rightarrow \|x_k - x^*\|^2 \leq \left(1 - \frac{2}{1 + L/M}\right)^{2k} \|x_0 - x^*\|^2$$

3.3) Our problem to solve is

$$\min \frac{1}{N}\sum f_i(w_i)$$
$$w_i = w_j \quad \forall j \in \mathcal{N}_i$$

Define $\bar{w}_{ik} = \frac{1}{[\mathcal{N}_i]} \sum w_{ij} \quad \forall j \in \mathcal{N}_i$

$$\Rightarrow g_i(\lambda) = L_i(w,\lambda) = f_i(w_{ik}) - \lambda^T(w_{ik} - \bar{w}_{ik})$$

if we define $A = I$ and $b_{ik} = \bar{w}_{ik}$

This is equivalent to

$$g_i(\lambda) = -f^*(-A\lambda) - \lambda^T b$$
$$= -f^*(-\lambda) - \lambda^T \bar{w}_{ik}$$

$\Rightarrow$ Dual Algorithm becomes

Step 1 : $w_{i,k+1} \in \underset{w}{\operatorname{argmin}}\, L(w_{ik}, \lambda_k) \ \forall i$

Step 2 : $\lambda_{k+1} = \lambda_k + \alpha_k (w_{ik+1} - \bar{w}_{ik}) \ \forall i$

Step 3 : $\overline{w_{ik}} = \frac{1}{[\mathcal{N}_i]} \sum w_{ij} \ \forall j \in \mathcal{N}_i$

~~Consensus~~

(Consensus)

The geometric graph converges at

$O(N \log N \log \epsilon^{-1})$

and we know the convergence of Dual ascent.

The combined convergence will be the worst of the two.