$$\text{prox}_t(\beta) = \underset{z}{\text{argmin}}\ \frac{1}{2t}\|\beta - z\|_2^2 + \|Dz\|_1$$

Dual: - conjugate loss f^*

- $\nabla f(\beta) + D^T u = 0$ solvable in β

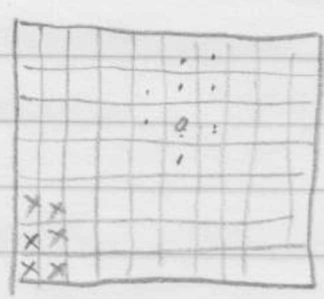
$$t f^*(-D^T u) + \phi(u)$$

$$t\cdot D \underbrace{\nabla^2 f^*(-D^T u)}_{\text{diag}} D^T + \underbrace{\nabla^2 \phi(u)}_{\text{diag}}$$

structured structured

FL over graph $D = \begin{bmatrix} -1 & \cdots & 1 \\ & -1 & 1 \\ -1 & \cdots & 1 \end{bmatrix}$

$D^T D = L$ graph Laplacian



0 or 1

no disease / disease

x_i y_i

$x_i^T \beta$ regularize according to $\sum_{\substack{(l,j) \\ \text{neighbors}}} |\beta_l - \beta_j|$

$$\min_{\beta} \; h(\underbrace{X\beta}_{=\alpha}) + \lambda \|\underbrace{D\beta}_{=z}\|_1$$

$$\Longleftrightarrow \min_{\beta,\alpha,z} \; h(\alpha) + \lambda \|z\|_1$$

$$\text{s.t.} \quad \underset{v}{X\beta = \alpha}, \quad \underset{u}{D\beta = z}$$

$$L(\beta,\alpha,z,u,v) = h(\alpha) + \lambda\|z\|_1 + u^T(D\beta - z) + v^T(X\beta - \alpha)$$

$$= h(\alpha) - v^T\alpha + \underbrace{\lambda\left(\|z\|_1 - \frac{u^T}{\lambda} z\right)} + \underbrace{(D^Tu + X^Tv)^T\beta}$$

$$\boxed{\begin{array}{lll} \text{Dual} & \min\limits_{u,v} & h^*(v) \\ & \text{s.t.} & D^Tu + X^Tv = 0 \\ & & \|u\|_\infty \le \lambda \end{array}}$$

- Dual prox grad

$$\text{prox}_t(u,v) = \underset{z,w}{\text{argmin}} \; \frac{1}{2t} \left\| \begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} z \\ w \end{pmatrix} \right\|_2^2$$

$$\text{s.t.} \quad D^Tz + X^Tw = 0$$

$$\|z\|_\infty \le \lambda$$

- Dual I.P. respect

$$D^Tu + X^Tv = 0.$$

$$A \begin{pmatrix} u \\ v \end{pmatrix} = 0$$

$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \; \end{bmatrix}$$

$$X^T \nabla h(X\beta) + D^T u = 0$$

E.g. Gaussian loss $h(\theta) = \frac{1}{2}\|y-\theta\|_2^2$

$$X^T(y - X\beta) = D^T u$$

$$X^T X \beta = X^T y - D^T u$$

solve a linear system in $X^T X$!

Primal s.g. still works, but slow

$$g = X^T \nabla h(X\beta) + \lambda D^T \gamma, \quad \gamma \in \partial\|D\beta\|_1$$

Primal prox grad

Reduces to solving

$$\text{prox}_t(\beta) = \underset{z}{\text{argmin}} \ \frac{1}{2t}\|\beta - z\|_2^2 + \lambda\|D\beta\|_1$$

solving this with a dual IP.

Free of linear systems in X!

$\min_u \ f^*(-D^T u)$ st $\|u\|_\infty \le \lambda$	$u + tD\nabla f^*(-D^T u)$ $\to$ project onto box