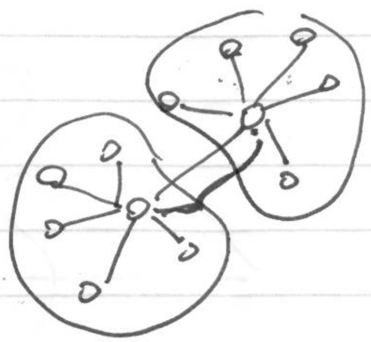
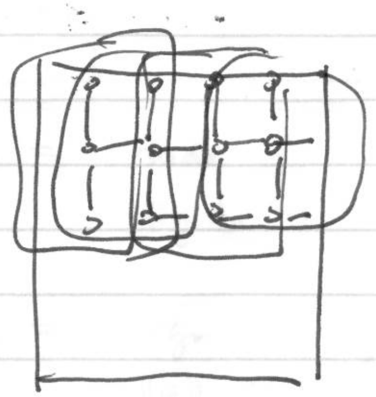




$$\min_{\beta} \; f(\beta) + \lambda \|D\beta\|_1 \quad \Big\} \quad \beta^+ = \beta - tg$$

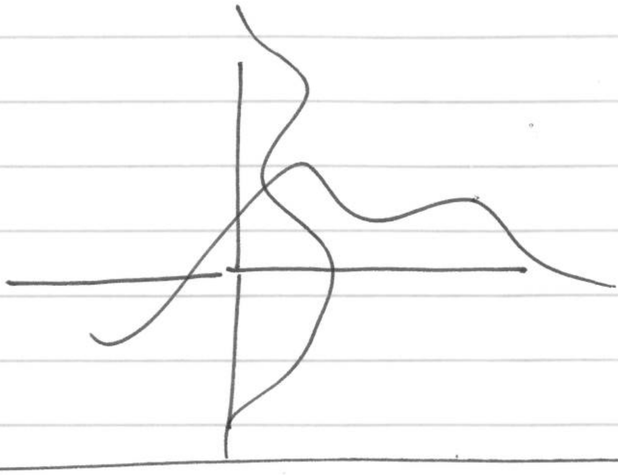
$$\nabla f(\beta) + \underline{\lambda D^T \gamma}, \text{ where } \gamma \in \partial \|\cdot\|_1 \Big|_{D\beta}$$

$$\nabla f(\beta) + \lambda \sum D_j \gamma_j$$

$$\gamma_j \begin{cases} = -1 & \text{if } (D\beta)_j < 0 \\ = 1 & (D\beta)_j > 0 \\ \in [-1, 1] & (D\beta)_j = 0 \end{cases}$$

valid choice of sub-grad is

$$\nabla f(\beta) + \lambda \sum_{j \in S} D_j \gamma_j, \quad S = \{ j : (D\beta)_j \neq 0 \}$$



$$\|x\|_p = \max_{\|y\|_q \le 1} y^T x \qquad \frac{1}{p}+\frac{1}{q}=1.$$

$$\partial\|x\|_p = \operatorname*{argmax}_{\|y\|_q \le 1} y^T x$$

$$\partial\|x\|_2 = \frac{x}{\|x\|_2} \qquad x \neq 0.$$

$$\in \{y : \|y\|_2 \le 1\} \qquad x = 0.$$

$$y_{11} \approx (\beta_1 y_{10} + \beta_2 y_9) + (\beta_3 y_8 + \beta_4 y_7) + \ldots (\beta_9 y_2 + \beta_{10} y_1)$$

$$= (\beta_1 X_{11,1} + \beta_2 X_{11,2}) + \ldots (\beta_9 X_{11,9} + \beta_{10} X_{11,10})$$

X matrix in row 11