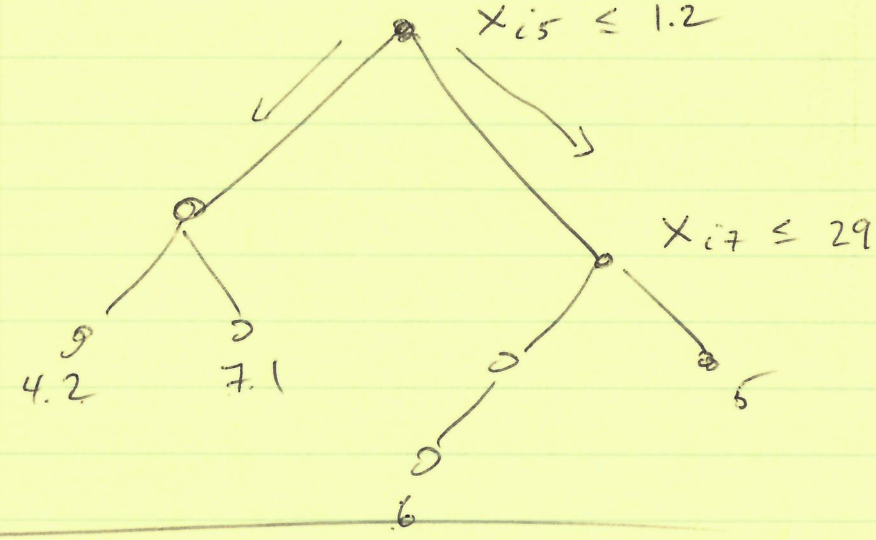
$x_i = (x_{i1}, \dots x_{ip})$



Find tree close to d

instead of ~~other~~

$u^{(k)} = u^{(k-1)} + t d$ (usual grad descent)

do $u^{(k)} = u^{(k-1)} + t \cdot \text{Tree}$

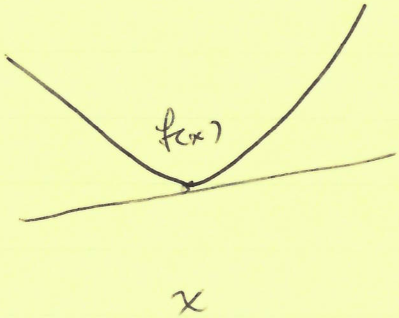
Hence find Tree that's "close to" d.

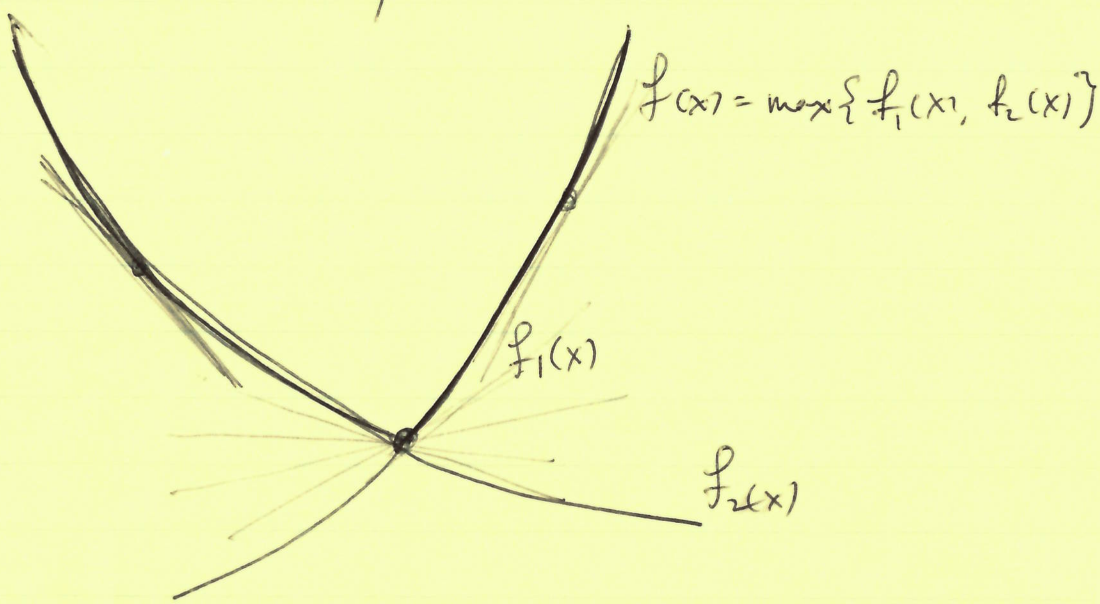
$\alpha_1 T_1 \quad + \qquad\qquad + \quad \alpha_r T_r$

SGD $\quad f(\beta) = \|y - X\beta\|_2^2$

$$= \sum_{i=1}^{n} \underbrace{(y_i - x_i^T \beta)^2}_{f_i(\beta)}$$

$$\boxed{g_{i_k}^{(k-1)} = \nabla f_{i_k}(x^{(k-1)})}$$





$g = \alpha \nabla f_1(x) + (1-\alpha) \nabla f_2(x)$

any $\alpha \in [0,1]$

$\partial f(x) = \{g : g \text{ is a sub-grad of } f \text{ at } x\}$

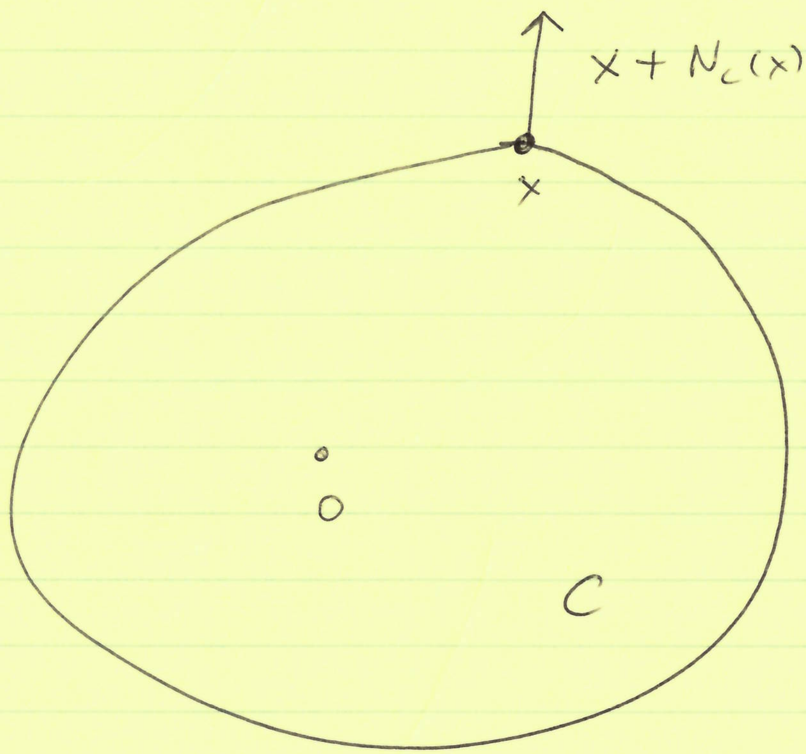
$g_1 \in \partial f(x), \quad g_2 \in \partial f(x)$

$(\alpha g_1 + (1-\alpha) g_2)^T (y - x) + f(x) \overset{?}{\leq} f(y)$ for all y

this is true b/c

$g_1^T (y-x) + f(x) \leq f(y) \quad \cdot \alpha$

$g_2^T (y-x) + f(x) \leq f(y) \quad \cdot (1-\alpha)$



$f(x) = I_C(x)$

$$\partial f(x) = \mathrm{conv}\Big(\bigcup_{i:\, f_i(x) = f(x)} \partial f_i(x) \Big)$$

$$f(x) = \max\{ f_1(x), f_2(x) \}$$

f_1, f_2 differentiable

$$\mathrm{conv}\big(\{ \nabla f_1(x), \nabla f_2(x) \}\big) = \{ \alpha \nabla f_1(x) + (1-\alpha) \nabla f_2(x) : \alpha \in [0,1] \}$$

$$\underbrace{\|x\|_p}_{f(x)} = \max_{z:\, \|z\|_q \le 1} \underbrace{z^T x}_{f_z(x)}$$