safe screening: if it passes, then we are guaranteed to have a 0 coefficient at solution.

non-safe rules: may know that w.h.p. have a 0 coefficient at solution or coefficient is small at sol. (means it has little statistical effect)

\[ \|x\|_1 = \max \frac{x^T z}{\|z\|_1} \] dual norm
\[ \|z\|_1 \leq 1 \]

\[ \partial \|x\|_1 = \text{argmax} \frac{x^T z}{\|z\|_1} \] subgradients.
\[ \|z\|_1 \leq 1 \]

\[ \|tz\| = t \|z\| \]

if we have \( \|x\|_\infty = \|x\|_1 \) (ith comp achieved the max)
then \( \text{sign}(x_i) \cdot e_i \in \partial \|x\|_\infty \) (ith spot \((0, \ldots, 1, \ldots, 0)\))
\[ f(s) \geq f(x^{(k-1)}) + \nabla f^T (s - x^{(k-1)}) \]

\[ \min_{s \in C} : \quad f^* \geq f(x^{(k-1)}) + \min_{s \in C} \nabla f^T (s - x^{(k-1)}) \]

Rearrange:

\[ -\min_{s \in C} \nabla f^T (s - x^{(k-1)}) \leq - f^* + f(x^{(k-1)}) \]

\[ \Rightarrow \quad f(y) = f(x) + \nabla f(x)^T (y - x) \]

\[ \Rightarrow \quad f(y) - f(x) - \nabla f(x)^T (y - x) \geq 0. \]

\[ \text{solve prob at } s \text{ if this is } \geq \varepsilon \]

\[ t_{k-1} \quad t_k \leftarrow \text{solve prob at } t_k \]