Convergence Analysis

Definition (Per-Node-Per-Iteration-Approximation Parameter)

At each iteration h, we define the accuracy level of the solution calculated by node k to its subproblem as

$$\theta_k^h := \frac{G_k^{\sigma'}(\Delta\alpha_k^{(h)}; v^{(h)}, \alpha_k^{(h)}) - G_k^{\sigma'}(\Delta\alpha_k^{\star}; v^{(h)}, \alpha_k^{(h)})}{G_k^{\sigma'}(\mathbf{0}; v^{(h)}, \alpha_k^{(h)}) - G_k^{\sigma'}(\Delta\alpha_k^{\star}; v^{(h)}, \alpha_k^{(h)})}$$

$\theta_k^h \in [0, 1]$, $\theta_k^h = 1$ means that no updates to the subproblem are made at iteration h

Assumption

Let $\mathscr{H}_h := (\alpha^{(h)}, \ldots, \alpha^{(1)})$ be the *dual vector history* until the beginning of iteration h, and define $\Theta_k^h := \mathbb{E}[\theta_k^h \,|\, \mathscr{H}_h]$. For all tasks k and all iterations h, we assume $p_k^h := \mathbb{P}(\theta_k^h = 1) \leq p_{max} < 1$ and $\hat{\Theta}_k^h = \mathbb{E}[\theta_k^h \,|\, \mathscr{H}_h, \theta_k^h < 1] \leq \Theta_{max} < 1$.

Convergence Analysis

Theorem 1

Assume that the losses f_k are $(1/\mu)$-smooth. Then, under Assumptions 1 and 2, there exists a constant $s \in [0, 1]$ such that for any given convergence target ε_D, choosing H such that

$$H \geq \frac{1}{(1 - \bar{\Theta})s} \log \frac{n}{\varepsilon_D}$$

will satisfy $\mathbb{E}[D(\alpha^{(H)}) - D(\alpha^{\star})] \leq \varepsilon_D$. Here, $\bar{\Theta} := p_{max} + (1 - p_{max})\Theta_{max} < 1$.