Federated Multi-Task Learning

Federated Update of W

Dual Problem

$$\min_{\alpha \in \mathbb{R}^N} \left[D(\alpha) := \sum_{k=1}^{K} \sum_{i \in \mathcal{P}_k} f_k^*(-\alpha_k^i) + \mathcal{R}^*(X\alpha) \right] \qquad X = diag(X_1, \ldots, X_K)$$

Data-local subproblem

$$\min_{\Delta\alpha_k \in \mathbb{R}^N} G_k^{\sigma'}(\Delta\alpha_k;\ v_k, \alpha_k) := \sum_{i \in \mathcal{P}_k} f_k^*(-\alpha_k^i - \Delta\alpha_k^i) + \langle w_k(\alpha), X_k \Delta\alpha_k \rangle + \frac{\alpha}{2} |\, X_k \Delta\alpha_k \,|_{M_k}^2 + c(\alpha)$$

Federated Multi-Task Learning, NIPS, 2017

Federated Multi-Task Learning

Federated Update of W

Algorithm 1 MOCHA: Federated Multi-Task Learning Framework

Input: Data $\mathbf{X}_t$ from $t = 1, \ldots, m$ tasks, stored on one of m nodes, and initial matrix $\boldsymbol{\Omega}_0$
Starting point $\boldsymbol{\alpha}^{(0)} := \mathbf{0} \in \mathbb{R}^n$, $\mathbf{v}^{(0)} := \mathbf{0} \in \mathbb{R}^b$
for iterations $i = 0, 1, \ldots$ **do**
    Set subproblem parameter σ' and number of federated iterations, H_i
    for iterations $h = 0, 1, \cdots, H_i$ **do**
        for tasks $t \in \{1, 2, \ldots, m\}$ **in parallel over** m **nodes do**
            call local solver, returning θ_t^h-approximate solution $\Delta\boldsymbol{\alpha}_t$ of the local subproblem (4)
            update local variables $\boldsymbol{\alpha}_t \leftarrow \boldsymbol{\alpha}_t + \Delta\boldsymbol{\alpha}_t$
            return updates $\Delta\mathbf{v}_t := \mathbf{X}_t\Delta\boldsymbol{\alpha}_t$
        reduce: $\mathbf{v}_t \leftarrow \mathbf{v}_t + \Delta\mathbf{v}_t$
    Update $\boldsymbol{\Omega}$ centrally based on $\mathbf{w}(\boldsymbol{\alpha})$ for latest $\boldsymbol{\alpha}$
Central node computes $\mathbf{w} = \mathbf{w}(\boldsymbol{\alpha})$ based on the lastest $\boldsymbol{\alpha}$
return: $\mathbf{W} := [\mathbf{w}_1, \ldots, \mathbf{w}_m]$