

パラメータ転移学習におけるリスク上界

理化学研究所 熊谷 亘

名古屋大学, 理化学研究所 金森敬文

本発表では, 機械学習における転移学習と呼ばれる問題設定を取り扱う. 転移学習の文脈では, 複数のドメイン (サンプル空間とその上の確率分布の組) を想定し, 目標のタスクが行われるドメインを目標ドメイン, その他のドメインを元ドメインと呼ぶ. 転移学習の目的は, 目標ドメインのサンプルに加えて元ドメインのサンプルを利用することで, 目標タスクに対してより良い性能を発揮するアルゴリズムを作ることである. 転移する知識の種類に応じて, 転移学習の問題へのアプローチはいくつかの種類に分類される. 本発表では, 各ドメインにある種のパラメトリックモデルが想定され, 転移された知識はパラメータにエンコードされるようなパラメータ転移アプローチについて考察する. 本研究の目的は, パラメータ転移アプローチに基づくアルゴリズムに対して, 理論的解析を行うことである.

本発表における目標ドメインおよび元ドメインにおけるモデルの定式化について述べる. 目標ドメインの仮説は, パラメトリックな特徴写像 $\psi_{\theta} : \mathcal{X}_{\mathcal{T}} \rightarrow \mathbb{R}^m$ を用いて, $h_{\mathcal{T}, \theta, \mathbf{w}}(\mathbf{x}) := \langle \mathbf{w}, \psi_{\theta}(\mathbf{x}) \rangle$ のような形で表されると仮定する. ここで θ と \mathbf{w} は仮説の持つパラメータである. また, 仮説 $h_{\mathcal{T}, \theta, \mathbf{w}}$ の期待リスクおよび経験リスクを $\mathcal{R}_{\mathcal{T}}(\theta, \mathbf{w})$ および $\hat{\mathcal{R}}_{\mathcal{T}}(\theta, \mathbf{w})$ のように記述する. 元ドメインには, サンプル分布や仮説などのパラメトリックモデルが存在するとし, パラメータ空間の一部 Θ は元ドメインと目標ドメインと共有されていると仮定する. そのとき, $\theta_S^* \in \Theta$ と $\mathbf{w}_S^* \in \mathcal{W}_S$ は元ドメインにおいて何らかの指標に関して有効なパラメータであるとする.

次に, 本研究で扱う転移学習アルゴリズムについて説明する. 元ドメインと目標ドメインでそれぞれ N 個と n 個のサンプルを使用できるとする. アルゴリズムは, まず N 個のサンプルを使用して, θ_S^* の推定値 $\hat{\theta}_N \in \Theta$ を出力する. アルゴリズムは次に目標ドメインのパラメータ

$$\mathbf{w}_{\mathcal{T}}^* := \operatorname{argmin}_{\mathbf{w} \in \mathcal{W}_{\mathcal{T}}} \mathcal{R}_{\mathcal{T}}(\theta_S^*, \mathbf{w})$$

に対し, n 個のサンプルを用いて推定値

$$\hat{\mathbf{w}}_{N,n} := \operatorname{argmin}_{\mathbf{w} \in \mathcal{W}_{\mathcal{T}}} \hat{\mathcal{R}}_{\mathcal{T},n}(\hat{\theta}_N, \mathbf{w}) + \rho r(\mathbf{w})$$

を出力する. ここで $r(\mathbf{w})$ は $\|\cdot\|_2$ に関して 1-強凸な正則化項とし, ρ は正の実数とする.

元ドメインが何らかの意味で目標ドメインに関係している場合, 元ドメインでの有効なパラメータ θ_S^* も目標タスクにとって有用であると期待される. 本発表では $\mathcal{R}_{\mathcal{T}}(\theta_S^*, \mathbf{w}_{\mathcal{T}}^*)$ を予測性能の基準値として採用する. このときいくつかの技術的仮定とともに, 本研究において新たに導入された局所安定性と転移学習可能性という条件を用いることで, 以下の学習バウンドを導出することができる (局所安定性および転移学習可能性の定義に関しては論文 [1] を参照せよ).

定理 1 (学習バウンド). パラメトリック特徴写像 ψ_{θ} が局所安定であるとする. また, 元ドメインで学習された $\theta_S^* \in \Theta$ の推定量 $\hat{\theta}_N$ は確率 $1 - \delta$ でパラメータ転移学習可能性を満たすとする. 正則化パラメータ ρ を適切に設定するとき, 以下の不等式が確率 $1 - (\delta + 2\delta)$ で漸近的に成り立つ:

$$\mathcal{R}_{\mathcal{T}}(\hat{\theta}_N, \hat{\mathbf{w}}_{N,n}) - \mathcal{R}_{\mathcal{T}}(\theta_S^*, \mathbf{w}_{\mathcal{T}}^*) \leq c \max \left\{ n^{-1/3}, \left\| \hat{\theta}_N - \theta_S^* \right\|^{2/7} \right\}$$

ここで c は N および n によらない定数である.

本研究に関連する他の理論的結果および数値実験の結果は当日報告する.

参考文献

[1] W. Kumagai, Learning Bound for Parameter Transfer Learning, *NIPS*, (2016).