2022年6月17日　統計数理研究所　オープンハウス

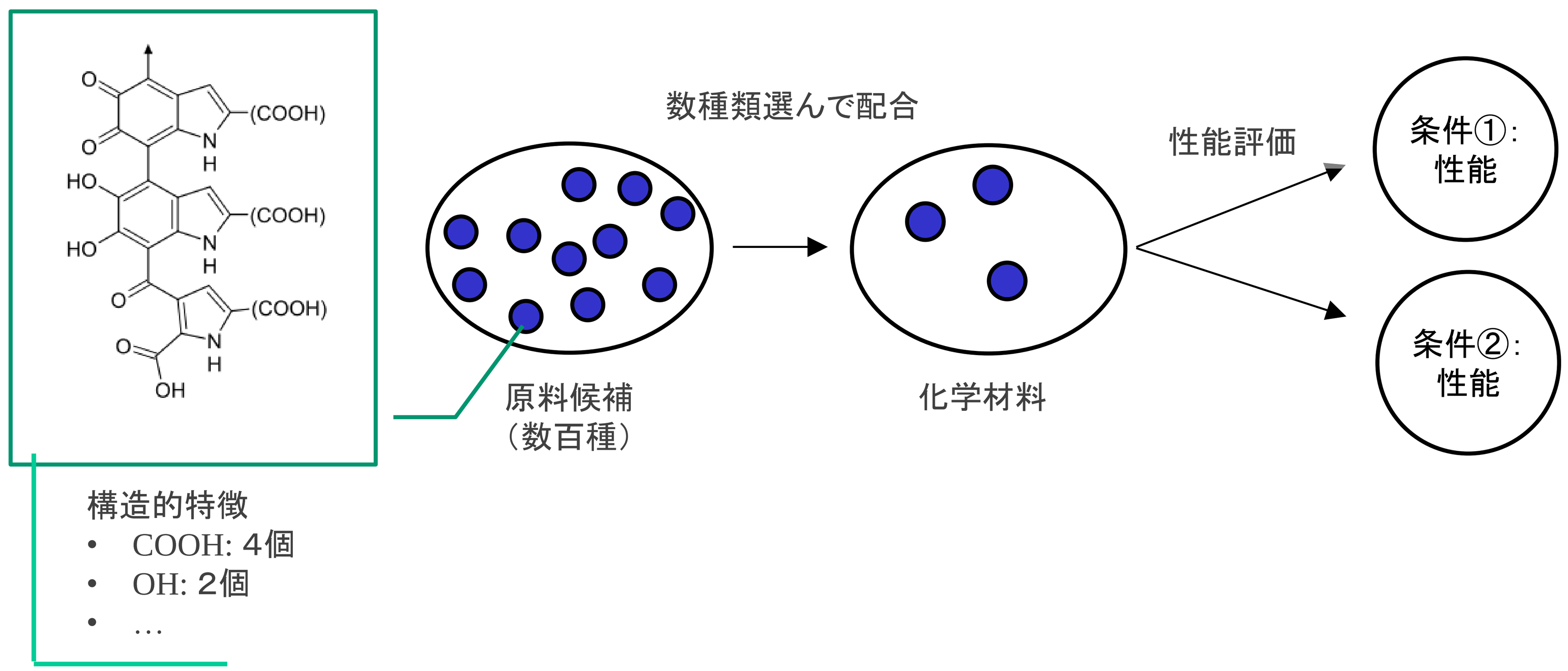
評価条件違いにおける階層モデルの転移学習

松尾　裕樹　　総合研究大学院大学統計科学専攻　博士課程4年

【背景】

化学企業では複数の原料を組み合わせることで化学材料が開発されている。
一般的に開発段階ではデータの取得コストが高い。

- データ数が少ない
- データの取得に人的コストと時間を要する



原料の使用量に対して線形回帰モデルによって性能を予測する際に、
原料の構造的特徴も考慮した階層モデルが有効である。

原料と性能の関係

$$y \sim Normal(W\beta, \sigma)$$

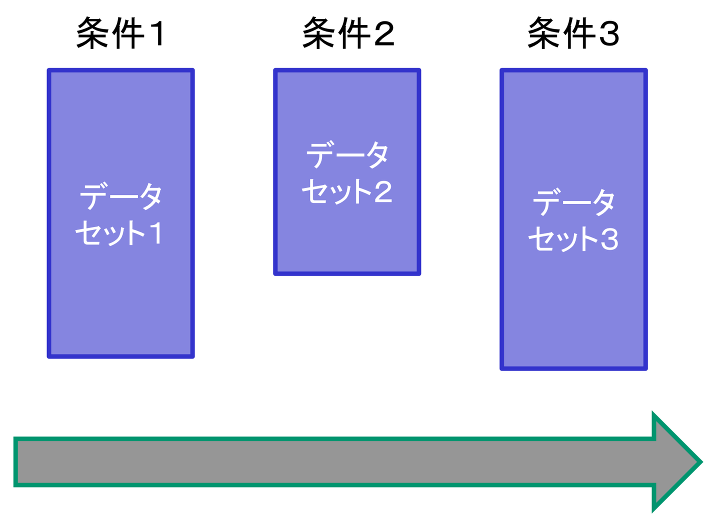
＋

構造的特徴の類似性

$$\Sigma = kernel_{rbf}(V, V \mid l)$$
$$\beta \sim MultiNormal(0, \Sigma)$$

⇩

$$\Sigma = kernel_{rbf}(V, V \mid l)$$
$$\beta \sim MultiNormal(0, \Sigma)$$
$$y \sim Normal(W\beta, \sigma)$$



データが蓄積しても条件が変わるごとにデータを取り直す必要がある

一方で、同じ化学材料、同じ性能であっても評価環境が変わることで
性能値が変化するため、評価条件が変わる度にデータを取り直す必要がある。

$$\Sigma_s = kernel_{rbf}(V_s, Vs \mid l) \qquad \Sigma_t = kernel_{rbf}(V_t, Vt \mid l)$$
$$\beta_s \sim MultiNormal(0, \Sigma_s) \qquad \beta_t \sim MultiNormal(0, \Sigma_t)$$
$$y_s \sim Normal(Ws\beta_s, \sigma) \qquad y_t \sim Normal(Wt\beta_t\, \sigma)$$

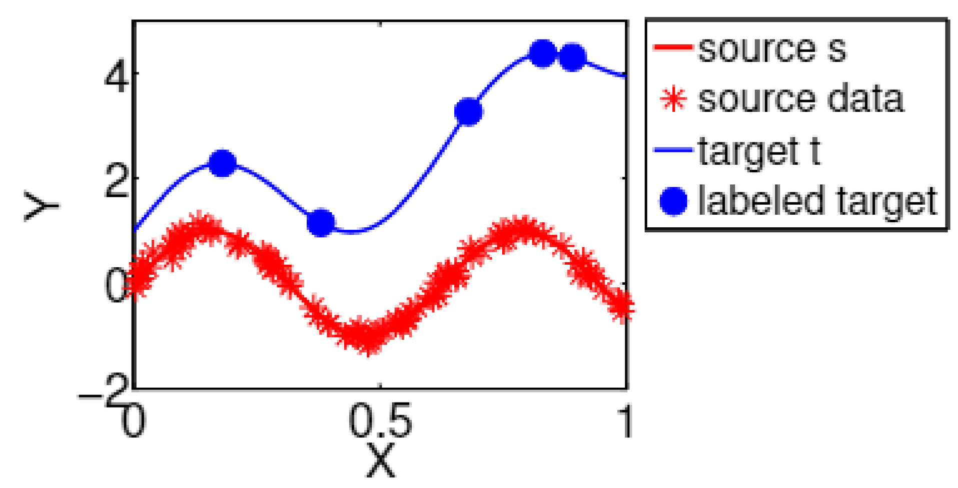
$$P(\beta_s|V_s) \neq P(\beta_t|V_t)$$

【Model Shift】

異なるドメイン間でモデル($P(Y|X)$)が変化する転移学習の問題設定の一つ。

Wang, X. and Schneider, J. G. Generalization bounds for transfer learning under model shift. In UAI, pp. 922–931, 2015.

$$P(Y_s|X_s) \neq P(Y_t|X_t)$$



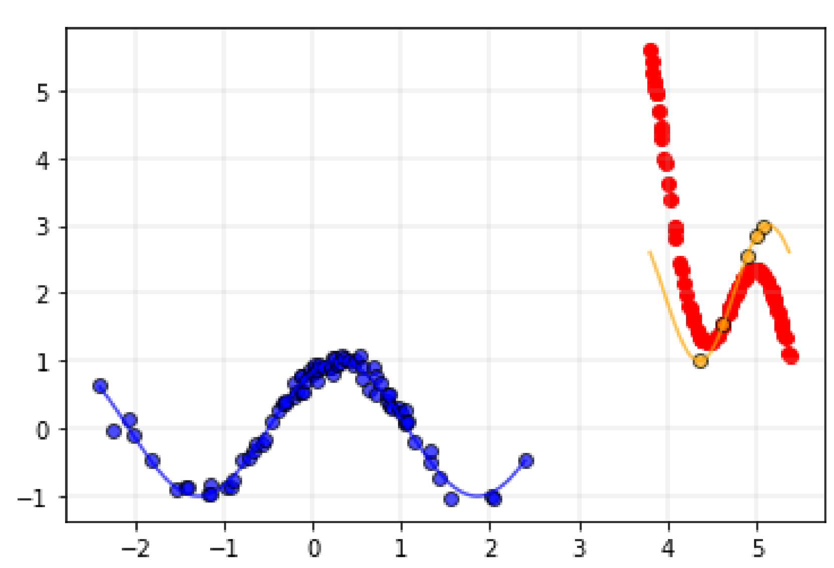
Model Shift の例.
同一の X に対してソースとターゲットでYが異なっている

【先行研究】

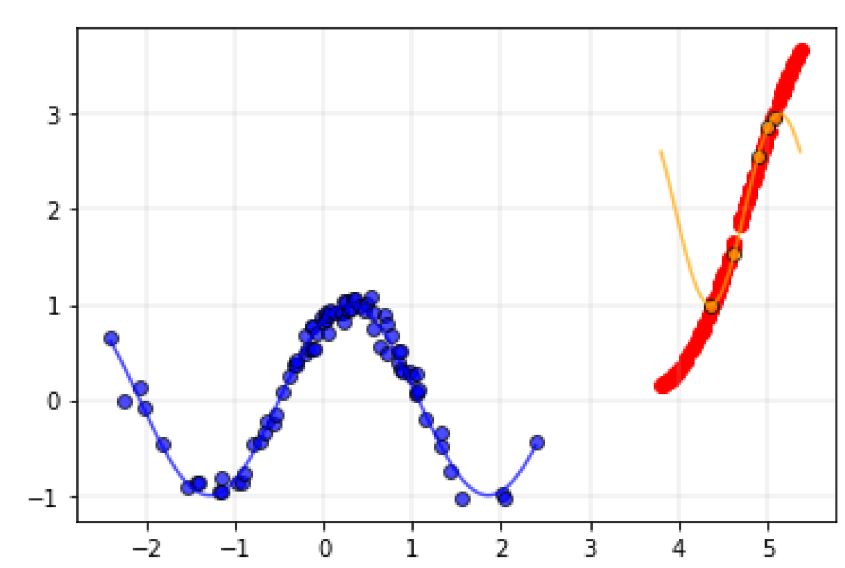
Support and Model Shift(SMS):
model shift に対応した転移学習手法。

1. ソースドメインでガウス過程回帰モデルを構築する。
2. ターゲットドメインデータをソースドメインへ変換し、ガウス過程回帰モデルで予測する。
3. 予測値をターゲットドメインへ再変換し、ターゲットドメインの予測値を得る。

Wang, X., Schneider, J.: Flexible transfer learning under support and model shift. In: NIPS. (2014)



SMSによるターゲットドメインの予測結果.
サインカーブを再現している.



GPRによるターゲットドメインの予測結果(target only) .
サインカーブを予測できていない。

【先行研究課題】

SMS中で使用されているガウス過程回帰は観測誤差が独立かつ等分散である必要があるため、
回帰係数の転移に適用できず、分散共分散を転移させることが出来ない。
分散共分散は以下の情報を有しており、ソースからターゲットへ転移させることで、
実験計画やベイズ最適化に活用できることが期待される。

- 学習データの偏り
- 変数間の共線性
- 各変数の推定精度

バッチベイズ最適化の獲得関数の一つである qEI
予測値の分散共分散を元に計算可能

$$qEI(X) = E_n\left[\left(f_n^* - \min_{i=1,\dots,q} f(x_i)\right)^+\right]$$

Matthias Schonlau. Computer experiments and global optimization. PhD thesis, Waterloo, Ont., Canada, Canada, 1997. AAINQ22234

分散共分散を転移させるモデル案

- Multi Task Gaussian Process Prediction
  タスク間に単純な相関関係を過程したガウス過程回帰モデル。

E. Bonilla, K.M. Chai and C. Williams, "Multi-Task Gaussian Process Prediction", Proc. 20th Ann. Conf. Neural Information Processing Systems, pp. 153-160, 2008.

$$\Sigma_v = kernel_{rbf}(V, V \mid l)$$
$$\Sigma_f = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$
$$S = \begin{pmatrix} \sigma_\beta^2 & 0 \\ 0 & \sigma_\beta^2 \end{pmatrix}$$
$$\begin{pmatrix} \beta_s \\ \beta_t \end{pmatrix} \sim MultiNormal\left(0, \Sigma_f \otimes \Sigma_v + S \otimes I_D\right)$$
$$y_s \sim Normal(W_s\beta_s, \sigma)$$
$$y_t \sim Normal(W_t\beta_t, \sigma)$$

【研究目的】

階層モデルにおいて回帰係数の分散共分散構造も含めて転移可能な手法を確立する。

【今後の予定】

分散共分散を転移可能な構造を事前分布に導入し、実装・検証を進める。