

シリサイド半導体のバンドギャップの線形回帰分析

Linear Regression Analysis of Band-Gap Energies of Semiconducting Silicides

九工大院情報工 淵 雅也, 有馬 幹尋, °前田 佳均

Dept. of Information Eng., Kyutech, Masaya Fuchi, Mikihiro Arima, °Yoshihito Maeda

E-mail: maeda@cse.kyutech.ac.jp

【はじめに】マテリアル・インフォマティクス(Materials Informatics: MI)は、物質・材料の物性・特性を既存データから統計的手法でモデル化し、それによって未知の物性・特性を推測する情報工学である [1]. MI では、ある物質母集合の特定の物性・特性 Y (目的変数) に相関する物理量: 記述子 $\{x_i\}$ を用いて最尤統計モデル $Y=F(\{x_i\})$ の構築を行う. 本研究では、16 種のシリサイド半導体のバンドギャップエネルギー (E_g) [2] を Elastic Net 正則化と一般化線形回帰 (GLM) を行い、最大尤度の E_g の統計モデルを検討した.

【解析方法】Elastic Net 正則化、一般化線形回帰 (GLM) に必要な R 言語スクリプトパッケージを CRAN ホームページ[3]から入手して編集し、解析に用いた.

【結 果】 目的変数 $Y=E_g$ の記述子 $\{x_i\}$ には、太陽電池材料カルコパーライト(ABC_2)のバンドギャップの推測に用いられた構成元素の電気陰性度(EN), 原子番号(AN), 融点(MP), 擬ポテンシャル半径 (PR), 周期律表の元素グループ番号(VA)を設定した[4]. 元素 1 と 2 の 2 元化合物では、記述子は全部で 10 個ある. ここで、第 1 元素: Fe, Mg, Ca, Ba, Ru, Re, Ir, Os, W, Cr, Mo, Mn, 第 2 元素: Si, (Ge)である. まず、過学習回避と目的変数への寄与の大きな記述子に縮約するために Elastic Net 正則化を行った. その結果、第 2 元素の EN2 と VA2 は除かれ、記述子のセットは、 $\{EN1, AN1, AN2, MP1, MP2, PR1, PR2, VA1\}$ に縮約された. この記述子セットを用い、相互作用項 $\{x_i x_j\}$ を考慮した GLM 解析の結果を Fig.1 に示す. 縦・横軸は標準化した E_g (sr E_g) であり、その統計誤差には正規分布を仮定した. 第 1 元素が同じの $OsSi$, Os_2Si_3 を除いて良好な線形回帰ができた. sr E_g の最尤統計モデルは、第 2 元素の MP2 と PR2 の回帰係数が零となり、ほぼ第 1 元素の記述子から成り、相互作用項を含む重回帰式で以下のように表せる.

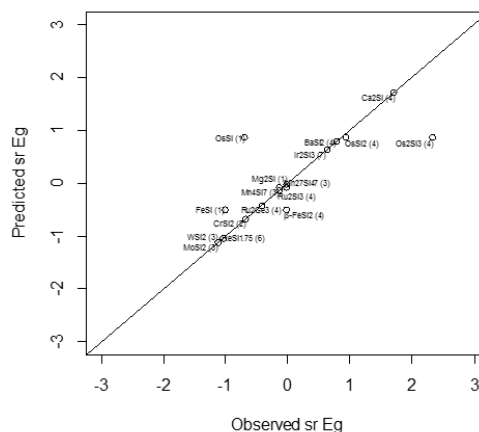


Fig.1 Result of linear regression using GLM for standardized band-gap energy E_g (sr E_g) of 16 semiconducting silicides.

$$\begin{aligned} \text{sr } E_g = & -1.500 - 1.212[AN1] - 0.067[AN2] + (1.483 + 3.252[AN1])[EN1] + (0.893 + 0.586[AN1] - 1.807[EN1])[MP1] \\ & + (1.302 + 1.435[AN1] - 1.549[EN1])[PR1] - (0.199 + 0.555[AN1])[VA1] \end{aligned}$$

【参考文献】 [1] K. Rajan ed.: *Informatics for Materials Science and Engineering*, Butterwoerh-Heinemann, 2013, Amsterdam [2] V. Borisenko: *Semiconducting Silicides*, Springer 2000, Berlin [3] CRAN Homepage, <https://cran.r-project.org/web/packages/MCMCpack/index.html> [4] P. Dey et al.: *Computational Materials Science* **83**, 185 (2014).