サンプル空間の対称性を利用した
線形回帰分析法
（記述子の対称化）

藤井将

目次

- 導入
- 4元ホイスラー合金の分類
- 4元ホイスラー合金の元素置換に対する対称性
- 積表、指標表、射影演算子
- 射影演算子による記述子の対称化
- 記述子の具体例
- 3元ホイスラー合金、ハイエントロピー合金への応用（検討中）
- 実際の計算（まだ検討中）

導入

線形回帰分析を行う際、
サンプル空間に対称性や、何らかの制限、要請がある場合、
それを事前知識として、どのように活用できるか検討する。

このようなことは、少なくともMI研究においては、
まだ誰も真剣に取り組んでいないと思われる。

対称性を見つけることの一つの利点は、
なんの苦労もすることなくサンプル物質を増やせることである。
（回帰においてサンプル物質は多いにこしたことはないので、
これは大きな利点である。）

対称性を見つけることのもう一つの利点は、
記述子の対称化である（後述）。
また、サンプル空間に何らかの制限・要請がある場合も、
その制限を記述子の制限に利用することができる可能性がある。
これらは無用な記述子を考えることを避けることができるばかりか、
物理的意味の抽出においても重要となる。

4元ホイスラー合金の分類

サイト1 (0, 0, 0) に元素 *A*
サイト2 (0.25, 0.25, 0.25) に元素 *B*
サイト3 (0.5, 0.5, 0.5) に元素 *C*
サイト3 (0.75, 0.75, 0.75) に元素 *D*
がいる4元ホイスラー合金を ***ABCD*** と表すことにする。

同じ組成*A*、*B*、*C*、*D*であっても、その並べ方を変えると
元の合金*ABCD*とは異なる結晶となる可能性がある。
並べ方は全部で 4! = 24 通りある。
これら24個の結晶は以下の3種類に分類される。

*ABCD*系:	*ABDC*系:	*ACBD*系:
ABCD	*ABCD*	*ACBD*
BCDA	*BDCA*	*CBDA*
CDAB	*DCAB*	*BDAC*
DABC	*CABD*	*DACB*
DCBA	*CDBA*	*DBCA*
CBAD	*DBAC*	*BCAD*
BADC	*BACD*	*CADB*
ADCB	*ACDB*	*ADBC*

*ABCD*系は*A*のNNが*B*、*D*、2nd N が*C*元素

*ABDC*系は*A*のNNが*B*、*C*、2nd N が*D*元素

*ACBD*系は*A*のNNが*C*、*D*、2nd N が*B*元素

（この対称性の発見により
サンプル空間を8倍にできる。）

4元ホイスラー合金の元素置換に対する対称性

以下、*ABCD*系に注目する。
A=1、*B*=2、*C*=3、*D*=4 とし、各等価な結晶（8個）を置換で表現する。
（行列表現しても良いが、簡単のため置換で表現。）

$$ABCD = 1234 = E\,1234$$
$$BCDA = 2341 = S_1\,1234$$
$$CDAB = 3412 = S_2\,1234$$
$$DABC = 4123 = S_3\,1234$$
$$DCBA = 4321 = S_4\,1234$$
$$CBAD = 3214 = S_5\,1234$$
$$BADC = 2143 = S_6\,1234$$
$$ADCB = 1432 = S_7\,1234$$

各置換操作の定義

これらは群を成す（対称群）。

$E = (1)(2)(3)(4)$ （恒等操作）
$S_1 = (1\ 2\ 3\ 4)$ （1つ左にスライドする操作）
$S_2 = (1\ 3)(2\ 4)$ （2つ左にスライドする操作）
$S_3 = (1\ 4\ 3\ 2)$ （3つ左にスライドする操作）
$S_4 = (1\ 4)(2\ 3)$ （順序反転操作）
$S_5 = (1\ 3)$ （順序反転+1つ左にスライドする操作）
$S_6 = (1\ 2)(3\ 4)$ （順序反転+2つ左にスライドする操作）
$S_7 = (2\ 4)$ （順序反転+3つ左にスライドする操作）

逆元

$E^{-1} = E$
$S_1^{-1} = S_3$
$S_2^{-1} = S_2$
$S_3^{-1} = S_1$
$S_4^{-1} = S_4$
$S_5^{-1} = S_5$
$S_6^{-1} = S_6$
$S_7^{-1} = S_7$

積表、指標表、射影演算子

積表

E	S_1	S_2	S_3	S_4	S_5	S_6	S_7
S_1	S_2	S_3	E	S_7	S_4	S_5	S_6
S_2	S_3	E	S_1	S_6	S_7	S_4	S_5
S_3	E	S_1	S_2	S_5	S_6	S_7	S_4
S_4	S_5	S_6	S_7	E	S_1	S_2	S_3
S_5	S_6	S_7	S_4	S_3	E	S_1	S_2
S_6	S_7	S_4	S_5	S_2	S_3	E	S_1
S_7	S_4	S_5	S_6	S_1	S_2	S_3	E

類: (E), (S_1, S_3), S_2, (S_4, S_6), (S_5, S_7)

指標表

	E	$2S_1$	S_2	$2S_4$	$2S_5$
Γ_1	1	1	1	1	1
Γ_2	1	1	1	-1	-1
Γ_3	1	-1	1	1	-1
Γ_4	1	-1	1	-1	1
Γ_5	2	0	-2	0	0

今の場合、興味があるのは全対称表現 Γ_1 のみ。
（ハイエントロピー合金の場合は反対称化された記述子が必要となるかもしれない）

全対称表現の射影演算子

$$
\begin{aligned}
P(\Gamma_1) &= E + S_1 + S_2 + S_3 + S_4 + S_5 + S_6 + S_7 \\
&= (E + S_4)(E + S_1 + S_2 + S_3)
\end{aligned}
$$

射影演算子による記述子の（全）対称化

1変数記述子の場合

$q(1)$: サイト1の元素種を変数に持つ任意の1変数記述子

1変数対称化記述子 $w^{(1)}$

$$w^{(1)}(1,2,3,4) = P(\Gamma_1)q(1) = (E + S_4)(E + S_1 + S_2 + S_3)q(1)$$
$$= \underline{q(1) + q(2) + q(3) + q(4)}$$

（係数は無視）

この対称化記述子$w^{(1)} = q(1) + q(2) + q(3) + q(4)$ は確かに
全ての対称操作に対し不変である。
しかし問題がある。
それは、*ABCD*系、*ABDC*系、*ACBD*系で全く同じ形であるということである。
つまり、この1変数対称化記述子だけではこれら3つの物質系を区別することはできない。
よって、最低でも2変数記述子が必要であることがわかる。
（これは対称性というより、サンプル空間から導かれる要請と言える。）

注意: この1変数対称化記述子が3つの物質系を区別できないからといって、
不必要であるというわけではない。
（他の元素組成（*ABFG* など）を持つ物質と区別するためには重要かも知れない。）

射影演算子による記述子の（全）対称化

2変数記述子の場合

$q(1,2)$: サイト1と2の元素種を変数に持つ任意の2変数記述子
$q(1,3)$: サイト1と3の元素種を変数に持つ任意の2変数記述子

2変数対称化記述子 $w^{(2)}$

$$
\begin{aligned}
w^{(2)}{}_1(1,2,3,4) &= P(\Gamma_1)q(1,2) \\
&= \underline{q(1,2) + q(2,3) + q(3,4) + q(4,1) + q(2,1) + q(3,2) + q(4,3) + q(1,4)}
\end{aligned}
$$

$$
\begin{aligned}
w^{(2)}{}_2(1,2,3,4) &= P(\Gamma_1)q(1,3) \\
&= \underline{q(1,3) + q(3,1) + q(2,4) + q(4,2)}
\end{aligned}
$$

2変数の場合、対称化記述子は上記のように2種類得られる。

射影演算子による記述子の（全）対称化

2変数記述子の場合

得られた2変数対称化記述子に具体的な元素を入れてみる。

$$w^{(2)}{}_1(1,2,3,4) = q(1,2) + q(2,1) + q(2,3) + q(3,2) + q(3,4) + q(4,3) + q(4,1) + q(1,4)$$

$$w^{(2)}{}_1(ABCD) = q(AB) + q(BA) + q(BC) + q(CB) + q(CD) + q(DC) + q(DA) + q(AD)$$

$$w^{(2)}{}_1(ABDC) = q(AB) + q(BA) + q(BD) + q(DB) + q(DC) + q(CD) + q(CA) + q(AC)$$

$$w^{(2)}{}_1(ACBD) = q(AC) + q(CA) + q(CB) + q(BC) + q(BD) + q(DB) + q(DA) + q(AD)$$

$$w^{(2)}{}_2(1,2,3,4) = q(1,3) + q(3,1) + q(2,4) + q(4,2)$$

$$w^{(2)}{}_2(ABCD) = q(AC) + q(CA) + q(BD) + q(DB)$$

$$w^{(2)}{}_2(ABDC) = q(AD) + q(DA) + q(BC) + q(CB)$$

$$w^{(2)}{}_2(ACBD) = q(AB) + q(BA) + q(CD) + q(DC)$$

それぞれ形が違う

よって、これらの記述子を使えば、*ABCD*系、*ABDC*系、*ACBD*系を区別できる。

記述子の具体例

$q(AB)$は2変数であれば何でも良い。
具体例としては、

- 1変数記述子の積 $q(AB) = q(A)q(B)$
- 1変数記述子の和や差の絶対値 $q(AB) = |q(A) + q(B)|$
- 1変数記述子の和や差の逆数 $q(AB) = 1/(q(A) + q(B))$
- 1変数記述子の和や差の絶対値の逆数 $q(AB) = 1/|q(A) + q(B)|$
- 本質的に2変数の記述子（*B*元素中の不純物*A*の磁気モーメント等）

など、色々考えられる。

（これらは後で対称化されるので、この段階で対称性を気にする必要はない。）

応用

3元ホイスラー合金への応用は簡単で、
$A = B$ とすればそのまま上の議論が使える。
（3元系では、$AACD$系（$X\alpha$構造）、$ACAD$系（$L2_1$構造）の2種類のみとなる。）

ハイエントロピー合金への応用する場合、
例えば、$A_{1-x}B_xCD$ という物質においては、
$A_{1-x}B_xCD = B_xA_{1-x}CD$ という自明な対称性があるので、記述子には、
$w(A,B,C,D,x) = w(B,A,C,D,1-x)$
という対称性が求められることになる。

この場合、記述子が $f(x)$ と $g(ABCD)$とに変数分離可能であると仮定すると、
$w(A,B,C,D,x) = f_S(x)g_S(ABCD)$ という、両方とも対称な記述子の他に、
$w(A,B,C,D,x) = f_A(x)g_A(ABCD)$ という、両方とも反対称な記述子も可能となる。

この場合は$g_A(ABCD)$という、元素の並び替え$(AB \rightarrow BA)$に対し
反対称な記述子も検討する必要がある。