様　式　C－19、F－19、Z－19　(共通)

科研費 KAKENHI

科学研究費助成事業　　研究成果報告書

平成 26 年　6 月 13 日現在

機関番号：32660
研究種目：基盤研究(C)
研究期間：2011～2013
課題番号：23510152
研究課題名（和文）遺伝子機能予測における数理工学指向アルゴリズムの開発

研究課題名（英文）Algorithms for prediction of MHC binders using mathematical optimization models

研究代表者
施　建明（Shi, Jianming）

東京理科大学・経営学部・教授

研究者番号：70287465

交付決定額（研究期間全体）：（直接経費）　2,400,000円、（間接経費）　720,000円

研究成果の概要（和文）：ワクチン開発に不可欠の主要組織適合抗原(MHC)クラスII におけるペプチド結合予測の中核はEntropyの最大化問題であり，その最大化問題を分数和計画の最適化問題に帰着できる．本研究では，予測に使われている数理モデルを分数和計画問題と定式化し，その最適解を見付けるアルゴリズムを提案することは特徴である．本研究では，このような分数和計画問題最適化問題に対して，線形緩和を用いてる２種類のアルゴリズムを提案した．計算機実験では下記の結果を得ている．1)計算時間は平均で既存方法の12% しかかからない．2)分数の個数が60であっても平均で，約14分(CPU time)で最適解を得られる．

研究成果の概要（英文）：Prediction of the binding ability of antigen peptides to major histocompatibility complex (MHC) class II molecules plays an important role in vaccine development. A key to such prediction is to maximize the entropy of information obtained from the the binding ability. This problem can be formulated as a maximization of the Sum of Ratios Problem. In this study, two kinds of algorithms which use linear relaxation were proposed to such a Sum of Linear Ratios Problem, that is an approximation of general Sum of Ratios Problem. The following results have been obtained in the numerical experiments.
1) The new algorithm takes a 12% CPU time of the existing methods on average. 2) The proposed algorithm finds an optimal solution of the Sum of Linear Ratios Problem with 60 ratios in about less than 14 minutes (CPU time).
These numerical results indicate that the proposed algorithms are much superior to theses existing algorithms.

研究分野：社会システム工学・安全システム

科研費の分科・細目：社会・安全システム科学・社会システム工学・安全システム

キーワード：分数計和計画問題　数理最適化　アルゴリズム　遺伝子機能予測　主要組織適合抗原(MHC)クラスII

1．研究開始当初の背景

医学の発達が著しい今日において，病気の治療手段も高度に発展されてきたと共に，医療費も年々増加傾向にある．発病の予防医療は様々な形があるが，ワクチンによる発病予防への期待が大きい．特に完治が困難ながんをはじめ，エイズ，薬剤久性菌感染症などの多くの病気に対してワクチンによる発病予防はもっとも低コストかつ有効手段の１つであると考えられる．

ワクチンの開発において主要組織適合遺伝子複合抗原(主要組織適合抗原，MHC)クラスIIにおけるペプチド結合予測が極めて重要な課題である．ペプチド結合予測において，様々なヒューリスティック方法が提案されているが，いずれも予測の精度と計算時間の効率性の両面で，質の高いものとは言い難い．一方，数理最適化の分野において予測の対象・目的を数式で記述することが可能ならば，数理最適化モデルとしての最適解を求めることは可能になる場合が多い．しかし，ペプチド結合予測を数理計画問題に帰着させると極めて難しい最適化問題になる．このような困難な最適化問題に対するアルゴリズムの開発への期待が大きい．

2．研究の目的

ペプチド結合予測において，ＯＲの観点から予測モデルを主に分数和計画の最適化問題として定式化し，効率のよい最適解を求めるアルゴリズムの開発は研究の目的である．

３．研究の方法

ペプチド結合予測に使われる数理モデルの中核は entropy の最大化問題である．この最大化問題を分数和最適化問題に帰着させる．

この非凸最適化問題に対して，線形緩和により，分数和最適化問題の最適解を求めるアルゴリズムを開発した．

4．研究成果

既存のアルゴリズムの計算時間が分数の個数の指数に比例するので，その個数が 20 ぐらいになると，現実的な時間で最適解を求めることは困難となる．その理由は，既存研究は目的関数の個々の分数を t/w にし，(t,w)空間に t/w の線形関数を緩和し，分枝限定アルゴリズムの設計ることにある．

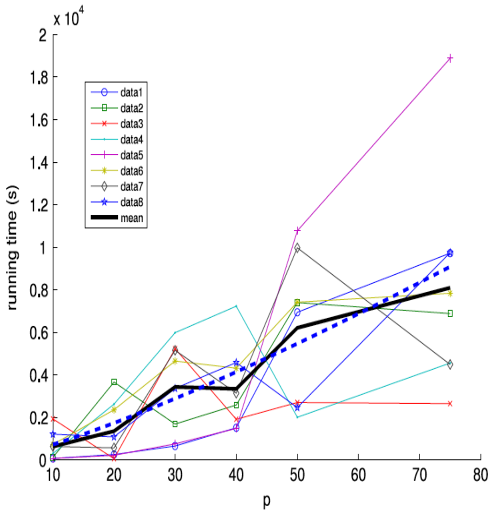
本研究の成果として，分数の和の数が変数の次元よりかなり多い場合，分枝限定の作業を変数の空間で行うことができるようアルゴリズムを設計した．これより，アルゴリズムの高速化の実現ができ，効率性の向上に繋がったことが，計算機実験により検証できた．

また，分数和最適化問題に対しては，下記のような等価な問題に帰着させる．さらに，分枝限定法を使用できるような線形緩和を作り，アルゴリズムを設計した．効率を測るために,計算機実験を共同研究者が Minnesota

$$
\begin{aligned}
\text{maximize} \quad & \sum_{i=1}^{p} \left(n_i^\top y^i + a_i z_i \right) \\
\text{subject to} \quad & d_i^\top y^i + b_i z_i = 1 && \forall i = 1, \ldots, p \\
& A y^i - c z_i \le 0 && \forall i = 1, \ldots, p \\
& \frac{1}{\beta_i} \le z_i \le \frac{1}{\alpha_i} && \forall i = 1, \ldots, p \\
& y^i z_j = y^j z_i && \forall i, j = 1, \ldots, p.
\end{aligned}
$$

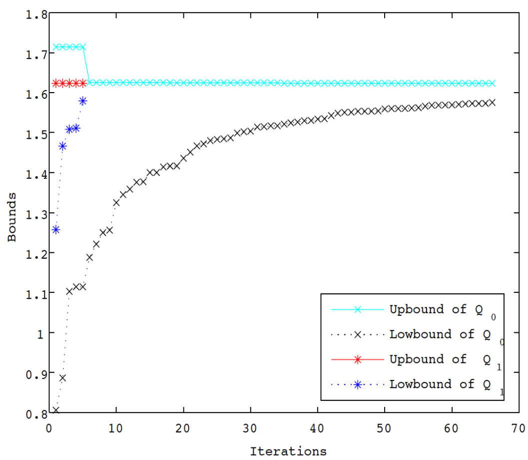
大学で行い，計算の効率性を大幅最善されることが確認された．

この実験に使われたデータはランダムで生成されたものである．下記の図では，分数の個数が 75 (p=75)までの実験結果を示してい



る．平均計算時間(太い黒線)がおおよそ p の線形比例になっていることも興味深い．つまり，既存のアルゴリズムが p の指数時間の実験結果に対して，本研究で開発したアルゴリズムは，実験結果として，p の線形時間で最適解を得ることができた．

また，本研究では，さらに強い線形緩和を試みた．この結果は〔雑誌論文〕①でまとめた．成果として分枝限定の特殊状況を考慮に入れ，$y^i z_j$ に対して既存の緩和より強い線形緩和を提案した．この方法により，計算の効



率性は著しく改善されたことが実験で確認された．計算機実験の結果は下記の図で示している．

この図では，変数が低次元(この実験では 3 次元)の問題に対して提案した方法 Q_1 を用いると，最適値の下界値と上界値の差が早い段階で小さくなり，最適解を見付けることができたことを示している．明らかに改善前の方法 Q_0 より早く収束していることがみえる．

方法 Q_1 に関して，その計算時間(CPU time)・反復回数・分枝の数の実験結果を下記の Table 1 でまとめている．どの指標をみても Q_1 の方が優れていることがわかる．

Model	# of p	CPU time(s)			# of Iterations			# of Branches		
		MIN.	AVERG.	MAX.	MIN.	AVERG.	MAX.	MIN.	AVERG.	MAX.
Q_1	2	0.13	0.25	0.54	1	1.71	2	1	1.67	2
Q_0	2	0.45	2.30	6.18	3	17.00	48	3	27.33	89
Q_1	5	0.97	6.12	11.85	1	2.80	5	1	2.16	5
Q_0	5	3.55	20.32	36.10	12	54.58	117	14	88.33	200
Q_1	10	16.90	52.78	203.20	5	9.00	24	7	18.00	54
Q_0	10	27.14	125.46	243.32	46	192.42	362	71	317.67	622
Q_1	15	33.38	84.63	173.05	5	12.80	28	7	21.00	47
Q_0	15	549.43	3684.40	12537.86	538	3610.80	12434	893	6882.40	23972

Table 1: Numerical results for solving problem (P_0) with Q_1 and Q_0 with various p and $n = 3$.

この計算機実験の結果によると，Q_1 の計算時間は平均で既存方法 Q_0 の 12% しかかからない．また，分数の個数が 60 であっても平均で,約 14 分(CPU time)で最適解を得ることに成功した．下記の Table 2 では p=60，次元が 3 である場合の計算機実験の結果を示している．新方法 Q_1 では，計算時間(CPU time)・反復回数・分枝の数のいずれも Q_0 より極めてよい結果を得ている．特に反復回数・分枝の数は極めて少ない．

Instance	Model	CPU time(s).	# of Iterations	# of Branches
I_1	Q_1	729.45	7	12
	Q_0	9162.30	4383	7424
I_2	Q_1	1278.31	12	18
	Q_0	27843.25	11166	18992
I_3	Q_1	436.94	5	9
	Q_0	1675.83	792	1142
I_4	Q_1	882.48	9	17
	Q_0	9066.27	3055	4957

Table 2: Results of the 4 instances of the numerical experiments with $p = 60$ and $n = 3$.

具体的な実験方法と環境や実験で得た成果の詳細は〔雑誌論文〕①をご参考頂く．

本研究では p は 15 と 60 の間において，研究期間の都合で，計算機実験を行っておらず，今後の課題になる．生なデータを使用し，ペプチド結合予測に関する性能評価の実験も行っておらず，生なデータを入手できた次第，実験を行う予定である．

5．主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計 4 件）

① Y.Hu, J.Shi and S.Watanabe, A Revised Algorithm for Solving the Sum of Linear Ratios Problem with Lower Dimension using Linear Relaxation, International Journal of Operations Research, 11, 2014, 28-39.

② T.Kitahara, S.Mizuno and J.Shi, The LP-Newton method for standard form linear programming problems, Operations Research Letters, 41, 2013, 426-429.

③ J.G.Carlsson and J.Shi, A linear relaxation algorithm with lower dimension for solving the sum of linear ratios problem, Operations Research Letters, 41, 2013, 381-389.

④ L.Gao, S.K.Mishra and J.Shi, An extension of branch-and-bound algorithm for solving sum-of-nonlinear-ratios problem Optimization Letters, 6, 2012, 221-230.

〔学会発表〕（計 3 件）

① Y.Hu and J.Shi, Algorithms for the sum of linear ratios program with lower dimension and related problems, The 9th International Conference on Optimization, Dec. 12-16, 2013.

② J.Shi, A computational geometric approach for solving linear programming: Toward strong polynomial, 21th International Symposium on Mathematical Programming, August 19-24, 2012. TU Berlin, Germany.

③ J.Shi, A new algorithm for linear programming, INFORMS, International Beijing 2012, June 24-27, 2012. China National Convention Center, Beijing, China.

〔図書〕（計 0 件）

〔産業財産権〕

○出願状況（計 0 件）

名称：なし
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

○取得状況（計 0 件）

名称：なし
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕
ホームページ等 なし

6．研究組織

(1)研究代表者
　　施　建明　　（SHI, Jianming）
　東京理科大学・経営学部・教授
　研究者番号：70287465

(2)研究分担者
　　なし