ショートノート

点集合の重み付きミニマックス線形近似問題に対するアルゴリズム†

今井 桂子†† 今井 浩†††

d 次元空間での n 点の重み付きミニマックス線形近似問題が，アルゴリズム的には $d+1$ 次元での $2n$ 点の凸包を求めるという計算幾何学での最も基本的な問題に帰着できることを示す．これより，$d=2$ の平面の場合には，問題が $O(n\log n)$ の最適の手間で解けることがわかる．

1. はじめに

d 次元空間において n 点を超平面で近似する問題は，統計・パタン認識・施設配置理論などでよく現れる基本的な問題である．点集合を超平面で近似するとき，その近似の基準は応用の仕方に応じてさまざまなものが考えられている．本稿では，各点に正の値の重みが与えられているとき，点と超平面との重み付き距離を（点と超平面の Euclid 距離）×（点の重み）と定め，点と超平面の重み付き距離の最大値を最小にするような近似超平面（重み付きミニマックス近似超平面と呼ぶ）を求める問題を考える．

$d=2$，すなわち平面におけるこの問題については，重みのない（すなわち，各点の重みが等しい）場合には $O(n\log n)$ の手間のアルゴリズム[1]が，重みのある場合には $O(n^2\log n)$ の手間のアルゴリズム[2]が与えられている．また，$d=2$ の場合の問題の計算複雑度が $\Omega(n\log n)$ であることも示されている[2]．$d=3$ での重みなしの問題に対しては，$O(n\log n+k)=O(n^2)$（k は n 点の3次元凸包の対せき点対の数を表す）の手間のアルゴリズムが与えられている[1]．

本稿では，d 次元での n 点に対するこの問題が，アルゴリズム的には $d+1$ 次元での $2n$ 点の凸包を求める問題に帰着できることを示す．凸包を求める問題は，計算幾何学での最も基本的な問題であり，有用な結果が得られている[3]．3次元での $2n$ 点の凸包は，$O(n\log n)$ の手間で求めることができるので，この結果より $d=2$ の重み付きの場合の問題が $\Theta(n\log n)$ の最適の手間で解けることがいえる．

2. 重み付きミニマックス線形近似問題

d 次元空間において，n 個の正の重み w_i の点 $p_i=x_i=(x_{i1},x_{i2},\cdots,x_{id})$ $(i=1,\cdots,n)$ が与えられている（$n>d$ とする）．この空間での超平面 $a^{\mathrm T}x+b=0$ $(a\neq 0)$ への点 p_i からの Euclid 距離は

$$|a^{\mathrm T}x_i+b|/\sqrt{a^{\mathrm T}a}$$

で表され，重み付きミニマックス近似問題は

$$\min_{a\neq 0,b}\max_{i=1,\dots,n} w_i|a^{\mathrm T}x_i+b|/\sqrt{a^{\mathrm T}a}$$

となる．d 次元の点集合に対し，点が同一の超平面上にないとき，その点集合は退化していないと呼ぶ．退化していない点集合の点数は，$d+1$ 以上である．

補題 1. 重み付きミニマックス近似超平面 $a^{*\mathrm T}x+b^*=0$ に対して，それからの重み付き距離が最大となる点の集合は，退化していない．

証明：重み付き距離が最大となる点の添字の集合を I とし，$p_i(i\in I)$ が同一の超平面 $a^{\mathrm T}x+b=0$ 上にあったとする．この a,b に対して，a^*,b^* をそれぞれ $a^*+\varepsilon a$, $b^*+\varepsilon b$ に更新することを考える．点 $p_i(i\in I)$ と超平面 $(a^*+\varepsilon a)^{\mathrm T}x+(b^*+\varepsilon b)=0$ との重み付き距離は，

$$w_i|(a^*+\varepsilon a)^{\mathrm T}x_i+(b^*+\varepsilon b)|/\sqrt{(a^*+\varepsilon a)^{\mathrm T}(a^*+\varepsilon a)}$$
$$=w_i|a^{*\mathrm T}x_i+b^*|/\sqrt{a^{*\mathrm T}a^*+2\varepsilon a^{\mathrm T}a^*+\varepsilon^2a^{\mathrm T}a}$$

である．したがって，ε を $a^{\mathrm T}a^*$ と同じ符号の十分小さな数としたとき，その ε に対する超平面の方が点 $p_i(i\in I)$ との重み付き距離の最大値の値が小さくなってしまう．これは $a^{*\mathrm T}x+b^*=0$ が重み付きミニマックス近似超平面であり，それから重み付き距離が最大

† An Algorithm for Weighted Minimax Linear Approximation of Points by KEIKO IMAI (Information Science Center, Kyushu Institute of Technology) and HIROSHI IMAI (Department of Computer Science and Communication Engineering, Kyushu University).

†† 九州工業大学情報科学センター
††† 九州大学工学部情報工学科

の点が $p_i(i\in I)$ であるということに反し，結論を得る．□

d 次元空間での各点 $p_i=(x_{i1},\cdots,x_{id})$ に対して，$d+1$ 次元目の軸 y を導入し，$d+1$ 次元空間での2点

$$q_i^+=(w_ix_{i1},\cdots,w_ix_{id},w_i)$$

$$q_i^-=-(w_ix_{i1},\cdots,w_ix_{id},w_i)$$

を考える．このようにして得られた $d+1$ 次元での $2n$ 点の集合を Q で，その $d+1$ 次元空間での凸包を CH(Q) で表す．以下，単に超平面といえば元の d 次元空間での超平面のことをいい，新たに導入された $d+1$ 次元空間における超平面は，d-超平面ということにする．すると，次の補題が得られる．

補題 2. 次の(ⅰ)，(ⅱ)は同値である．

(ⅰ) d 次元空間において，$a^Tx+b=0$，$a^Ta=1$ なる超平面で，次の条件をみたすものが存在する．超平面との重み付き距離が最大となる点の集合 $\{p_i|i\in I\}$ が退化しておらず，この超平面は点集合 $\{p_i|i\in I\}$ を $\{p_i|i\in I^+\}$ と $\{p_i|i\in I^-\}$ とに二分する（$|I^+|,|I^-|>0$）．

(ⅱ) $d+1$ 次元空間において，凸包 CH(Q) の面で，$q_i^+(i\in I^+)$，$q_i^-(i\in I^-)$ のみで張られ，$y=0$ と交わり，それを含む d-超平面で $a^Tx+by+c=0$，$a^Ta=1$ と表せるものが存在する．

証明：“(ⅰ)→(ⅱ)”(ⅰ)の条件より

$$w_i|a^Tx_i+b|=D \quad (i\in I) \tag{1}$$

$$w_j|a^Tx_j+b|<D \quad (j\notin I) \tag{2}$$

である．(1)式に関しては，

$$a^T(w_ix_i)+bw_i \quad (i\in I^+),$$

$$a^T(-w_ix_i)+b(-w_i) \quad (i\in I^-)$$

の $|I|$ 個の値はすべて同じ値で，D か $-D$ である．(2)式より

$$-D<\pm\{a^T(w_jx_j)+bw_j\}<D \quad (j\notin I)$$

である．したがって，$i\in I^+$ に対して，

$$a^T(x-w_ix_i)+b(y-w_i)=a^Tx+by+c=0$$

（ただし，$c=-w_ia^Tx_i-bw_i$ とする）は凸包の支持 d-超平面である．$p_i(i\in I)$ は d 次元で退化していないことより，この支持 d-超平面と凸包との交わりは面である．この面が $y=0$ と交わるのは自明．

“(ⅱ)→(ⅰ)” 上の証明の逆をたどればよい．□

3. アルゴリズム

前章の議論より，次のアルゴリズムが得られる．

1. 点 p_i に対して，q_i^+，q_i^- を考えその凸包 CH(Q) を求める；
2. 凸包 CH(Q) の $y=0$ と交わる各面 f_j に対して，その面を含む d-超平面 $a_j^Tx+b_jy+c_j=0$，$a_j^Ta_j=1$ を求め，その面上の1点を q_i^+ としたとき $d_j=|a_j^T(w_ix_i)+b_jw_i|$ を求める（d_j は q_i^+ のとり方によらないで面 f_j だけで決まる）；
3. $\min\{d_j\}$ を達成する d_k を求める；このとき，重み付きミニマックス近似超平面は $a_k^Tx+b_k=0$ である．

このアルゴリズムの手間は，ステップ1で $d+1$ 次元空間で $2n$ 個の点の凸包を求めるところで抑えられる（ステップ2でかかる手間は，凸包の面数に比例するものであり，それだけの手間は必ずステップ1ですでにかかっている）．この $d+1$ 次元の凸包は，$d=2$ のとき $O(n\log n)$ の手間で，$d\geqq3$ のとき $O(n^{\lceil(d+1)/2\rceil})$ の手間で求めることができる[3]（求める凸包の面数に手間が依存するアルゴリズムも与えられている[4]）．以上より次の定理を得る．

定理 1. d 次元での n 点に対する重み付きミニマックス線形近似問題は，$d=2$ のとき $O(n\log n)$ の最適の手間で，$d\geqq3$ のとき $O(n^{\lceil(d+1)/2\rceil})$ の手間で解ける．□

4. おわりに

本稿では，点集合の線形近似問題で，与えられた点と超平面との重み付き Euclid 距離の最大値を最小にするような超平面を求める問題を考えた．統計などでは，他に Euclid 距離の二乗の総和を最小にする直交最小二乗法や Euclid 距離の総和を最小にする直交 L_1 法[5] などが用いられる．アルゴリズムの観点からは，古くから最小二乗法に対するものが確立されていた．本稿では，ミニマックス法，すなわち L_∞ 法に対する効率のよい解法を与えた．

謝辞 本研究の一部は，統計数理研究所グループ共同研究（63-共研-13）の援助を受けた．

参考文献

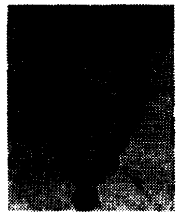
1) Houle, M. E. and Toussaint, G. T.: Computing the Width of a Set, *Proc. 1st ACM Symp. on Computational Geometry*, Baltimore MD, pp. 1-7 (1985).
2) Lee, D. T. and Wu, Y. F.: Geometric Complexity of Some Location Problems, *Algorithmica*, Vol. 1, pp. 193-211 (1986).
3) Preparata, F. P. and Shamos, M. I.: *Computational Geometry: An Introduction*, Springer-Verlag, New York (1985).

4) Seidel, R.: Output-Size Sensitive Geometric Algorithms, Ph. D. Thesis, Department of Computer Science, Cornell University (1986).
5) Yamamoto, P., Kato, K., Imai, K. and Imai, H.: Algorithms for Vertical and Orthogonal L_1 Linear Approximation of Points, *Proc. 4 th ACM Symp. on Computational Geometry*, Urbana, IL, pp. 352-361 (1988).

(昭和63年9月12日受付)
(平成元年2月14日採録)

今井　桂子（正会員）

1958年生．1980年津田塾大学学芸学部数学科卒業．1982年同大学院理学研究科専攻修士課程修了．1985年同大学院理学研究科博士課程単位取得退学．東京大学工学部計数工学科教務職員，助手を経て，1988年4月より九州工業大学情報科学センターに助手として勤務．現在，計算幾何学の分野で研究を行っている．



今井　浩（正会員）

昭和33年生．昭和56年東京大学工学部計数工学科卒業．昭和61年同大学院工学系研究科情報工学専門課程修了．工学博士．同年九州大学工学部情報工学科助教授となり，現在に至る．その間，昭和62年1月より4月まで，カナダマッギル大学計算機科学科訪問副教授．アルゴリズム全般，特に計算幾何学，組合せ最適化の研究に興味をもつ．電子情報通信学会，日本オペレーションズリサーチ学会，ACM 各会員．