

混合エキスパートモデルによる人工衛星軌道情報の予測

Prediction of Satellite Orbit Information by Mixture of Experts Model

菅原 啓介[†]
Keisuke Sugawara

新田 猛[†]
Takeru Nitta

石濱 直樹[†]
Naoki Ishihama

1. はじめに

宇宙開発が進むにつれて、多くの人工衛星が地球周回軌道にて運用されるようになってきている。しかし軌道上には衛星だけでなく、衛星の打上に用いられたロケットの残骸や、故障して廃棄された衛星、衛星同士の衝突によって発生した部品なども存在している。これらはスペースデブリと呼ばれ、運用中の衛星や国際宇宙ステーションと衝突することにより甚大な被害をもたらす可能性がある。衝突を未然に防ぐためには、衛星とデブリの軌道を正確に把握し、もし将来的に衝突する確率が高いことが予測される場合には適切な回避運用（推進力を発生させて軌道を変えること）を実施する必要がある。

米国の Combined Space Operations Center (CSpOC) では、地球を周回している衛星やデブリ等を観測して軌道を推定し、その情報を協力関係にある機関に Conjunction Data Messages (CDM) として数時間おきに配信している [1]。宇宙航空研究開発機構 (JAXA) でも、CDM に記載された軌道情報をもとに JAXA が運用する衛星とデブリ等が衝突する確率を算出してリスクを評価し、回避運用を検討している。類似した例として台風が接近するというシナリオを考えると、気象予報士が天気予報を配信し、行政機関がリスク評価や住民の避難計画を検討することに例えることができる。天気予報では、例えばある地域に 4 日後に台風が上陸する確率が予報円として示されるが、時間経過とともに予報が次第に変わっていき、結局上陸しなかったということが起こりうる。このようなケースが多いとリスク評価にかかる人的・時間的コストが大きくなるため、できる限り頻度を減らすことが必要である。CDM も同様に、あくまでも数日後の予報値が CSpOC から定期的に配信されるため、最接近の 4 日前では衝突する確率が高いと予想されていたものの次第にその確率が小さくなり、最終的に衝突することはなかったというケースも多い。スペースデブリの個数は増加傾向にあるため、接近事象が高頻度で起こるようになり、衛星運用者が衝突リスク評価にかかるコストが増大することが懸念されている。

したがって本研究では、接近事象の中から真に高リスクな事象を識別するために、将来得られる CDM を予測することを目的とする。比較的単純なモデルで将来 CDM を予測する研究は既に存在するが、本研究では CDM の特徴を十分にとらえられるような、より複雑なモデルによる予測を行う。提案する予測手法によって真に高リスクな接近事象の識別が可能となり、運用者の負荷軽減が期待される。

本論文の構成は次の通りである。まず第 2 章で CDM によるリスク評価方法について述べ、将来 CDM 予測の問題設定を行う。第 3 章では関連研究を紹介し、第 4 章では提案手法について述べる。第 5 章では実験結果と考察について述べ、最後に第 6 章でまとめを行う。

2. 問題設定

2.1 準備：衝突リスクの評価方法

JAXA では、軌道上の 2 物体の衝突リスク評価のために、2 物体の大きさや座標と速度から衝突確率を算出するという方法を採用している [2]。それらの値は CDM に記載されたものを使用しているため、今後の CDM の値の変化を予測することは、今後の衝突リスクの変化を予測することに等しい。

軌道上の物体は地上に設置されたレーダなどによって観測されるが、測定値には誤差が含まれているため、その座標推定値はある程度の不確実性をもつ。したがって CDM では、衛星の座標は点としてではなく確率分布として扱われている。より具体的には、衛星の 3 次元座標は、平均を表す 3 次元ベクトルとその分散を表す 3×3 の分散共分散行列によって表現される。

接近する 2 物体の位置関係と衝突確率計算の簡単なイメージについて図 1 に示す。直感的には、2 物体の座標平均が近く、座標共分散の重なりが大きいほど衝突確率が大きくなる。

以上のことから、ある時刻以降の衝突確率の変化を正確に予測するには、その時刻までに得られた CDM を用いて今後得られる CDM を正確に予測することが必要となる。表 1 に示す通り CDM には 2 物体の座標や速度以外の情報も記載されている。CDM の特徴を分析した結果、例えば図 2 に示すように大気抵抗を受けにくい種類の衛星の場合は座標の分散が小さくなりやすかったり、あるいは最接近 4 日前の時点で座標共分散が比較的大きい接近ケースは最接近直前になっても共分散が大きいままであったりと、現在の CDM と未来の CDM では相関があることが示唆されているため、予測が可能であると考えられる。

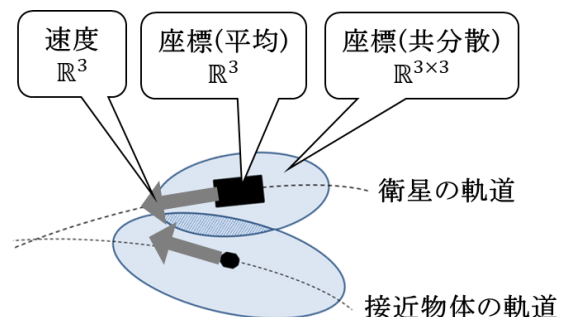


図 1. 軌道上 2 物体の衝突確率計算のイメージ

[†]宇宙航空研究開発機構, Japan Aerospace Exploration Agency.

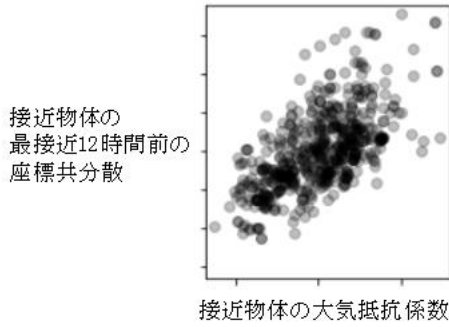


図 2. 最接近 4 日前と 12 時間前の CDM の関係

表 1. CDM に記載される変数の一例。

例えば最接近時の衛星の R 座標平均は SAT1_R となり、接近物体の表面積に関する値は SAT2_AREA_PC となる。また衝突確率の計算に必要な変数には◎をつけている。

CDM に含まれる変数名	意味
SAT{1,2}_{R,T,N}	◎最接近時の座標平均
SAT{1,2}_{COV_L}{R,T,N}_LOG	◎最接近時の座標の共分散の大きさを表す値
SAT{1,2}_{DELTA}{R,T,N}_DOT	◎最接近時の速度
SAT{1,2}_{CD_AREA_OVER_MASS}	大气抵抗係数
SAT{1,2}_{AREA_PC}	表面積に関する値

2.2 定式化

前述の通り、本研究は 2 物体が最接近する日時の 4 日前時点で得た CDM から、最接近 12 時間前に得られる CDM を予測することを目的とする。また CDM の値を点推定するのではなく、確率分布として推定しサンプリングが行えるようにしたい。したがって最接近 4 日前の CDM の値を $\mathbf{x} \in \mathbb{R}^M$ 、最接近 12 時間前の CDM の値を $\mathbf{y} \in \mathbb{R}^D$ とすると、この問題は条件付き確率 $p(\mathbf{y}|\mathbf{x})$ を推定する問題として定式化することができる。ただし最接近 4 日前の CDM に含まれる変数の中から、将来 CDM の予測に寄与するもののみを選択し $M = 34$ とする。また最接近 12 時間前の CDM については、衝突確率の計算に必要な 24 変数のみを予測するため $D = 24$ とする。

3. 関連研究

3.1 将来 CDM 予測

衝突リスク評価用ソフトウェア CAESAR[3]では、最接近直前の CDM の変数のうち座標共分散を正規分布としてモデル化し、衝突リスクの評価に用いている。

Hejduk らの手法[4]は、2 物体の共分散と大きさの確率密度分布を推定し、サンプリングを行って衝突リスクを評価する。

上記の 2 手法は、2 物体の速度については予測していない。Coppola[5]は速度の不確実性は衝突確率に影響を与えることを示している。このことから確率計算の入力となる速度も予測対象とし正確に予測することで、衝突確率をより正確に予測することができるようになると思われる。ただし速度データは多峰性を持つことが示唆されており、単純な正規分布ではその性質を表現することができないため、既存研究より表現能力の高いモデルを使用する必要がある。

3.2 生成モデル

本研究では、CDM の不確実性も含めてモデル化し、またサンプリングも行いたいため正規分布による生成モデルを採用する。また CDM データは多峰性を持つことが示唆されているため、複数の正規分布から構成される混合モデル[6]が適している。また本問題では条件付き確率 $p(\mathbf{y}|\mathbf{x})$ を推定する必要があるため、混合モデルの中でも条件付き確率を推定する混合エキスパートモデルを用いる[7]。

4. 提案手法

4.1 アプローチ

まず提案手法の大まかな方針について述べる。図 3(a)に示すように \mathbf{x}, \mathbf{y} 空間に散らばっている訓練データ点から尤もらしい条件付き確率 $p(\mathbf{y}|\mathbf{x})$ を推定する、というのが本研究の問題設定である。訓練データにない点 \mathbf{x}_{new} が入力されたとき、対応する確率変数 \mathbf{y}_{new} は、訓練データの中で \mathbf{x} 座標が \mathbf{x}_{new} と近い点 (図 3(b)中では明るい色で色付けした点) の \mathbf{y} 座標と近い値を取る可能性が高いと考えるのが自然である。また \mathbf{y}_{new} の値はそれらの点の \mathbf{y} 座標に正規分布に従うノイズ (図 3(b)では各訓練データ点のひげ) が乗った値を取ると仮定する。そのようにして、 \mathbf{x}_{new} に対して \mathbf{y}_{new} の確率密度関数を図に示すように生成することができる。

したがって、訓練データ集合をエキスパートとし、エキスパートを \mathbf{x}_{new} との距離で重みづけて、ある分散を持つ正規分布を重ね合わせることで \mathbf{y}_{new} が生成されるということができる。

また見方を変えると、 \mathbf{x} の値によって重みづけされた \mathbf{y} のカーネル密度推定と考えることもできる。

以上のことから、訓練データをフィッティングするモデルは「エキスパートと \mathbf{x}_{new} の距離尺度」と「エキスパートに固定された正規分布の分散」という 2 つのパラメータを持つ。次節でモデルを厳密に定式化し、この 2 パラメータの最適化方法について述べる。

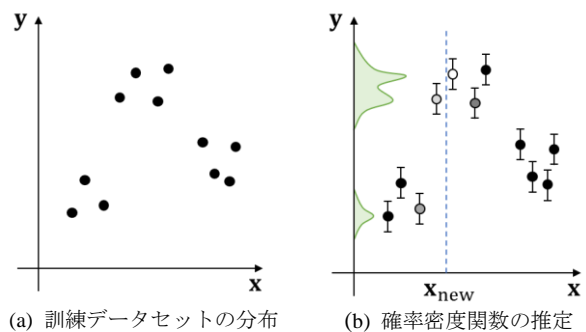


図 3. 提案手法の方針

4.2 モデルの定義とフィッティング方法

訓練データ集合 $\mathcal{D} = \{(\mathbf{x}_1, \mathbf{y}_1), \dots, (\mathbf{x}_N, \mathbf{y}_N)\}$ を次の混合エキスパートモデルにフィッティングさせ、 \mathbf{x} に関する \mathbf{y} の条件付き確率密度関数を推定する。

$$p(\mathbf{y}|\mathbf{x}; \mathbf{S}, \Sigma) = \sum_{k=1}^N \pi_k(\mathbf{x}; \mathbf{S}) \mathcal{N}(\mathbf{y}|\mathbf{y}_k, \Sigma)$$

混合係数 $\pi_k(\mathbf{x}; \mathbf{S})$ は以下に示す通り入力ベクトル \mathbf{x} の活性 $a_k(\mathbf{x}; \mathbf{S})$ の Softmax 関数として表されるゲート関数である。

$$\pi_k(\mathbf{x}; \mathbf{S}) = \frac{\exp(a_k(\mathbf{x}; \mathbf{S}))}{\sum_{j=1}^N \exp(a_j(\mathbf{x}; \mathbf{S}))}$$

ここで $a_k(\mathbf{x}; \mathbf{S})$ は正定値対称行列 $\mathbf{S} \in \mathbb{R}^{M \times M}$ からなる、ベクトル $\mathbf{x}_k \in \mathbb{R}^M$ とのマハラノビス距離の2乗にマイナスをかけた値である。

$$a_k(\mathbf{x}; \mathbf{S}) = -(\mathbf{x} - \mathbf{x}_k)^T \mathbf{S}^{-1} (\mathbf{x} - \mathbf{x}_k)$$

データ集合 \mathcal{D} に対する負の対数尤度 L は次の式で示される。

$$L = - \sum_{n=1}^N \ln \left\{ \sum_{k=1}^N \pi_k(\mathbf{x}_n; \mathbf{S}) \mathcal{N}(\mathbf{y}_n | \mathbf{y}_k, \mathbf{\Sigma}) \right\}$$

また \mathbf{S} および $\mathbf{\Sigma}$ の固有値が極端に小さくなることを防ぐために、行列式の対数にマイナスをかけた値を正則化項 R とする。各エキスパートが自分自身（すなわち $n = k$ となる点）の尤度のみを高くする方向に最適化が進むのを防ぐ効果が期待される。

$$R = -\lambda_1 \ln(|\mathbf{S}|) - \frac{\lambda_2}{2} \ln(|\mathbf{\Sigma}|)$$

ただし正の実数 λ_1, λ_2 は正則化の強さを決定づけるハイパーパラメータである。

以上から、最小化する目的関数 E は次の通りとなる。この目的関数 E をEMアルゴリズムによって最適化する。

$$E = L + R$$

Eステップでは、モデルパラメータ $\mathbf{S}, \mathbf{\Sigma}$ を固定し次式によって負担率 $\gamma_{n,k}$ が算出される。

$$\gamma_{n,k} = \frac{\pi_k(\mathbf{x}_n; \mathbf{S}) \mathcal{N}(\mathbf{y}_n | \mathbf{y}_k, \mathbf{\Sigma})}{\sum_{j=1}^N \pi_j(\mathbf{x}_n; \mathbf{S}) \mathcal{N}(\mathbf{y}_n | \mathbf{y}_j, \mathbf{\Sigma})}$$

Mステップではモデルパラメータ \mathbf{S} および $\mathbf{\Sigma}$ を更新する。 \mathbf{S}^{-1} についての偏微分は次の通りとなる。

$$\frac{\partial E}{\partial \mathbf{S}^{-1}} = - \sum_{n=1}^N \sum_{k=1}^N \gamma_{n,k} \left\{ \mathbf{X}_{n,k} - \sum_{j=1}^N \pi_j(\mathbf{x}_n; \mathbf{S}) \mathbf{X}_{n,j} \right\} + \lambda_1 \mathbf{S}$$

ただし $\mathbf{X}_{n,k} = (\mathbf{x}_n - \mathbf{x}_k)(\mathbf{x}_n - \mathbf{x}_k)^T$ とされている。この式を用いて勾配法によって最適な \mathbf{S}^{-1} を求める。

また $\mathbf{\Sigma}$ についての偏微分は次の通りとなる。

$$\frac{\partial E}{\partial \mathbf{\Sigma}} = \frac{1}{2} \mathbf{\Sigma}^{-1} \sum_{n=1}^N \sum_{k=1}^N \gamma_{n,k} \{ \mathbf{I} - \mathbf{Y}_{n,k} \mathbf{\Sigma}^{-1} \} - \frac{\lambda_2}{2} \mathbf{\Sigma}^{-1}$$

ただし $\mathbf{Y}_{n,k} = (\mathbf{y}_n - \mathbf{y}_k)(\mathbf{y}_n - \mathbf{y}_k)^T$ とされている。 $\frac{\partial E}{\partial \mathbf{\Sigma}} = 0$ となる次式の $\mathbf{\Sigma}_{\text{opt}}$ にそれぞれ更新する。

$$\mathbf{\Sigma}_{\text{opt}} = \frac{\sum_{n=1}^N \sum_{k=1}^N \gamma_{n,k} \mathbf{Y}_{n,k}}{N - \lambda_2}$$

直感的には、正則化の強さを決める λ_2 の値が大きいと $\mathbf{\Sigma}_{\text{opt}}$ の各要素の値が大きくなるといえる。

以上に示したE,Mステップを繰り返し、最適なモデルのパラメータ $\mathbf{S}, \mathbf{\Sigma}$ を求める。

5. 実験

JAXAで運用されている衛星Aに軌道上物体が接近した事象に関するCDMを704件集め、各変数が平均0、分散1になるように正規化を行った。また前述したモデルを訓練データ576件にフィッティングし、残りの128件をテストデータとして \mathbf{x} をモデルに入力して事後確率密度関数を求め、それぞれ5件ずつのサンプリングを行った。

5.1節では \mathbf{x} から \mathbf{y} の予測が可能であることを確かめ、5.2節ではサンプリングされた点の \mathbf{y} の分布について既存研究との比較を行う。

5.1 条件付き確率密度関数の推定

図4では $\mathbf{x} = [x_1, \dots, x_{34}]^T$ の要素のうち接近物体の気抵抗の受けやすさを示す x_{30} を横軸に、また $\mathbf{y} = [y_1, \dots, y_{24}]^T$ の要素のうち接近物体の座標の共分散の大きさを示す y_{18} を縦軸にとり、訓練データ点とテストデータ点、およびフィッティング済み混合エキスパートモデルからサンプリングした点を可視化している。訓練データ点から、気抵抗係数が大きい物体は座標の共分散の大きさも大きくなる傾向を持つことが読み取れるが、提案手法によってその傾向をモデル化できていると言える。

また別の変数にも実験を行った。図5では接近物体の表面積に関する要素 x_{28} を横軸に、接近物体の座標の共分散の大きさを示す要素 y_{20} を縦軸に取っている。概ね \mathbf{x} と \mathbf{y} の関係はモデル化できていると言える。また λ_1 の値を変えた場合のフィッティング結果の違いについても確かめた。4.0, 1.0, 0.50の3種類の値を試行し、 $\lambda_1 = 1.0$ のときに最も負の対数尤度が小さくなった。

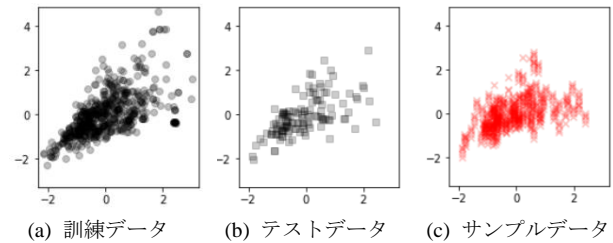


図4. 最接近4日前の気抵抗係数と最接近12時間前の座標の共分散の関係について可視化した結果。
横軸： x_{30} (最接近4日前のSAT2_CR_AREA_OVER_MASS)
縦軸： y_{18} (最接近12時間前のSAT2_COV_LR_LOG)

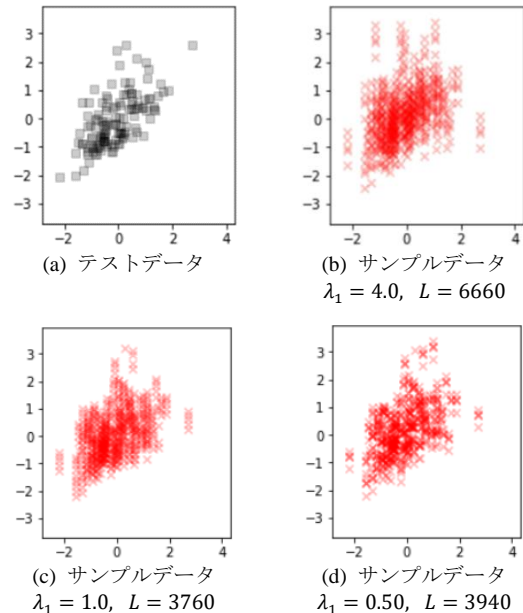


図5. パラメータ λ_1 によるサンプル点と負の対数尤度 L の違いについての評価結果。

横軸： x_{28} (SAT2_AREA_PC)
縦軸： y_{20} (最接近12時間前のSAT2_COV_LN_LOG)

5.2 既存研究との比較

CAESAR[3]では、最接近直前の CDM の変数のうち座標共分散を単一の正規分布にフィッティングさせている。比較のため、訓練データセットを正規分布にもフィッティングさせて得られた確率密度関数からサンプリングを行って \mathbf{y} の値を得た。

まず接近物体の座標共分散結果を図 6 に示す。 \mathbf{y} から接近物体の座標共分散に関する 2 要素 y_{18}, y_{20} を選んで縦軸・横軸にとり、(a) から (c) の各図に示した。

訓練データを見ると、第 1 象限と第 3 象限とは分散が異なっていることがわかることから、 \mathbf{y} に関する局所的な構造があることが示唆される。提案手法によって混合エキスパートモデルにフィッティング・サンプリングした結果 (b) と CAESAR の手法に基づいて単一正規分布にフィッティング・サンプリングした結果 (c) について比較すると、CAESAR の手法では局所的な構造が表現できていないが、提案手法では局所的な構造を比較的表現できていくことがわかる。

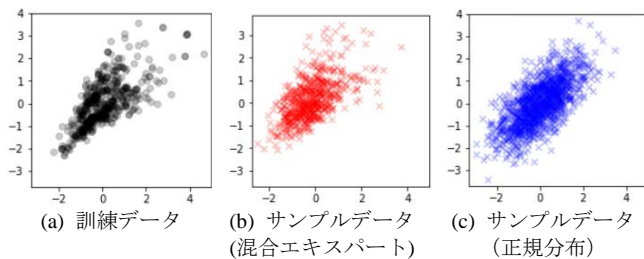


図 6. 接近物体の座標共分散の比較

横軸： y_{18} (最接近 12 時間前の SAT2_COV_LR_LOG)

縦軸： y_{20} (最接近 12 時間前の SAT2_COV_LN_LOG)

また、フィッティング済み混合エキスパートモデルと正規分布のテストデータに関する負の対数尤度を求めると、次の通り正規分布の方が低い値となった。

- ・混合エキスパートモデル：3012.3
- ・正規分布：1225.8

図 7 に示す通り、提案した混合エキスパートモデルは一部のデータに対して負の対数尤度が極端に大きくなる傾向が読み取れる。図 8 に示す負の対数尤度の累積和を見ると、混合エキスパートモデルの値は正規分布混合の値を 100 件目付近で超えることがわかる。よって提案したモデルは、訓練データにも含まれる典型的なデータに対する尤度は高くなるが、外れ値に対しては低いため総合的に負の対数尤度関数値は小さくなっていくことが示唆される。

したがって、特にハイパーパラメータ λ_2 の値をうまく選び、外れ値もカバーできるような Σ を得ることが重要である。

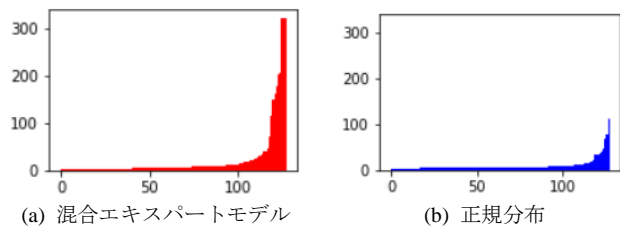


図 7. テストデータ 128 件について、縦軸に負の対数尤度を取り、値が小さい順に左から並べた図。

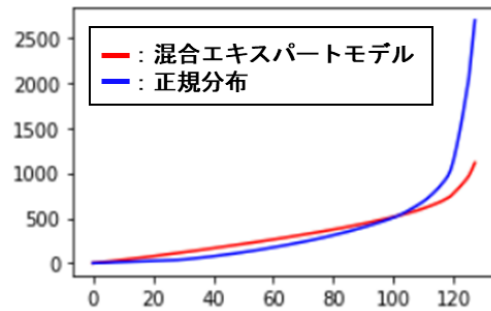


図 8. 図 7 で示した各データの累積和

また、Hejduk らの研究[4]ではモデリング対象外だった衛星と接近物体の速度を表す 2 要素の値の分布を図 9 に示す。十字型の多峰性を持つ分布を正しく推定できていることがわかる。

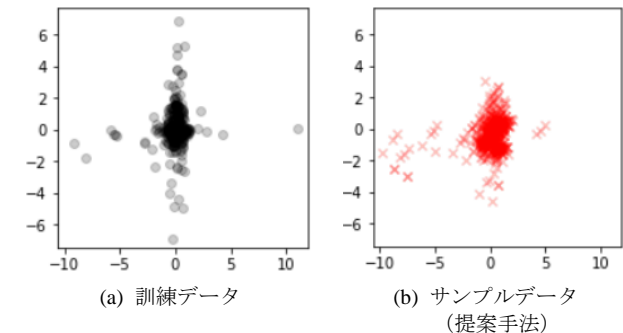


図 9. 衛星 A と接近物体の速度

横軸： y_4 (最接近 12 時間前の SAT1_DELTA_T_DOT)

縦軸： y_{16} (最接近 12 時間前の SAT2_DELTA_T_DOT)

6. おわりに

人工衛星とスペースデブリ等の衝突リスクを早期に評価するために、混合エキスパートモデルによって軌道情報を予測する手法を提案した。実際の接近情報を用いて実験を行い、軌道情報のサンプル点を生成できることを示した。また既存研究では予測できていない変数も予測できることを示した。一方、ハイパーパラメータのチューニングや、外れ値への対処に課題が残った。より多くのハイパーパラメータの探索や計算の効率化を行い、精度の高いモデルを得られるようにしたい。

参考文献

- [1] *Conjunction Data Message (Blue Book)* (Issue brief No. CCDS 508.0-B-1, Washington, DC, 2013), pp. 1-72.
- [2] Klinkrad, H. *Space Debris – Models and Risk Analysis*, Praxis Publishing, Chichester, UK, 2006
- [3] Laporte, F. “JAC Software, Solving Conjunction Assessment Issues.” AMOS Technical Conference, Maui HI, 2014
- [4] Hejduk, M. D., & L. C. Johnson, “Approaches to evaluating probability of collision uncertainty.” AAS Space Flight Mechanics Meeting, Napa, CA, 2016.
- [5] Coppola, V. T. “Including Velocity Uncertainty in the Probability of Collision Between Space Objects”, *Proceedings of the 22nd AAS/AIAA Space Flight Mechanics Meeting*, Vol. 143, Advances in the Astronautical Sciences, Univelt, San Diego, CA, 2012, pp. 2159–2178
- [6] Bishop, C. *Pattern Recognition and Machine Learning*, Springer, Berlin, Heidelberg, pp. 430-439, 2006.
- [7] Jacobs, R. A., M. I. Jordan, S. J. Nowlan, and G. E. Hinton “Adaptive mixtures of local experts.” *Neural Computation* 3(1), pp.79–87, 1991.