

ロバスト主成分分析に基づく沖合津波観測網の最適設計 (シンポジウム講演概要)

Optimal Design for Configuration of Offshore Tsunami Gauges with Robust Principal Component Analysis (Proceedings of Symposium on Applied Mechanics)

藤田 真粹 (東北大学) 野村 怜佳 (東北大学) 大竹 雄 (東北大学)
越村 俊一 (東北大学) 森口 周二 (東北大学) 寺田 賢二郎 (東北大学)

Saneiki FUJITA, Tohoku University

Reika NOMURA, Tohoku University

Yu OTAKE, Tohoku University

Shunichi KOSHIMURA, Tohoku University

Shuji MORIGUCHI, Tohoku University

Kenjiro TERADA, Tohoku University

E-mail: saneiki.fujita.s8@dc.tohoku.ac.jp

To realize real-time tsunami risk prediction, we present an optimization scheme to configure offshore wave gauges using machine learning and sparse sensing techniques. The present approach employs robust principal component analysis to extract low-dimensional spatial mode from a dataset of pre-computed tsunami wave heights. The sparse gauges' configuration is optimized according to the impact of observation effects only by the low-rank spatial modes that are supposed to dominate the gauges far from shore. As a demonstration, the optimal observation network off the coast of Shikoku is designed for 1564 possible Nankai Trough earthquakes. The forecast capability of the detected gauges' configuration is examined within the framework of the previous tsunami scenario detection.

1. 緒言

近年、津波被害低減を目的に、沖合での津波観測データを利用したリアルタイムリスク予測手法が提案されている(例えば1)。このような予測システムを国土全体で実現するには、広範囲にわたって密なリアルタイム観測網を導入する必要があるが、計測器の設置や運用に要する費用が問題となる。そのため、襲来する津波の情報を効率的に観測することが可能なスパース観測網の設計が必要となる。ただし、沖合を伝播する波高に対して、沿岸部における波高が相対的に大きくなるという津波の特徴と、リアルタイムリスク予測に求められる「即時性」との二つを踏まえると、絶対値の小さな沖合での伝播成分をより重用するような組み合わせ最適化問題から観測網を決定する必要がある。

本研究ではロバスト主成分分析(RPCA: Robust Principal Component Analysis)とスパースセンサーの最適配置手法²⁾を用いて津波観測網の最適設計法を提案し、南海トラフ地震・津波を想定した四国沿岸域を対象として計測点の最適配置を特定する。そして、そこでの観測情報を用いた津波シナリオ同定に基づく即時予測法¹⁾を実施し、予測精度の観点から本手法の有用性について検証する。

2. 手法

2.1. 観測点配置の最適化

本研究では、 n 点の候補位置で任意時刻の波高情報を復元可能な p ($p < n$) 点の組み合わせを特定することを観測点配置の最適化と呼ぶ。具体的には、ある時刻における候補点での波高 $\mathbf{x}_t \in \mathbb{R}^n$ と、観測可能な p 点での波高情報 $\mathbf{y}_t \in \mathbb{R}^p$ がダウンサンプリング行列 $\mathbf{C} \in \mathbb{R}^{p \times n}$ を用いて次のように関連付ける。

$$\mathbf{y} = \mathbf{C}\mathbf{x} = \mathbf{C}\Phi_r \mathbf{a} \quad (1)$$

ここで、 $\Phi_r \in \mathbb{R}^{n \times r}$ 、 $\mathbf{a} \in \mathbb{R}^r$ はそれぞれ、主成分分析によって抽出・次元削減された空間モード、係数ベクトルであり、

空間モード Φ_r を介することで n 次元の波高情報を r 次元まで縮約した情報を有する。この r 次元の状態量は、次式のように観測情報 $\mathbf{y}$ から逆推定することが可能である。

$$\hat{\mathbf{a}} = (\mathbf{C}\Phi_r)^\dagger \mathbf{y} \quad (2)$$

ここで、 $(\bullet)^\dagger$ は $\bullet$ の一般化逆行列である。 $\mathbf{a} \approx \hat{\mathbf{a}}$ となる観測点配置を特定することにより、スパース観測情報から本来の n 次元空間情報 $\mathbf{x}$ の推定を可能にすることが最適化の目的となる。具体的には、次式のように係数の誤差共分散行列の行列式を最小化する行列 $\mathbf{C}$ を特定することで最適化を行う。

$$\arg \min_{\mathbf{C}} \det [\text{Var}(\mathbf{a} - \hat{\mathbf{a}})] = \arg \min_{\mathbf{C}} \det \left[(\mathbf{C}\Phi_r)^\top \mathbf{C}\Phi_r \right]^{-1} \quad (3)$$

この組み合わせ最適化問題を効率的に解くために、列ピボット付き QR 分解に基づく貪欲法²⁾を適用する。

$$(\Phi_r \Phi_r^\top) \mathbf{C}^\top = \mathbf{Q}\mathbf{R} \quad (4)$$

ここで、 $\mathbf{Q}$ は直交行列であり、 $\mathbf{R}$ は上三角行列である。また、 $\mathbf{C}$ は $\mathbf{R}$ の対角成分を降順に並び替えるための置換行列とみなされる。

2.2. ロバスト主成分分析

主成分分析(PCA: Principal Component Analysis)は教師なし機械学習手法の一種であり、データ行列の主要な特徴量抽出と次元縮約を可能とする。しかし、データの分散を参照する PCA は外れ値の存在に敏感であるという欠点がある。すなわち、津波波高データへ PCA を適用する場合、抽出される特徴量は波高の大きな沿岸部観測点の情報に強く依存してしまう。

そこで本研究では、RPCA を適用することで沿岸部での情報を効果的に除去して沖合での伝播挙動に対して支配的な特徴量の抽出を目指す。RPCA では、対象とするデータ

行列 X が、外れ値を格納したスパース行列 S と秩序構造を格納した低ランク行列 L に分解できるものとし、以下の最適化式を解くことにより各行列 L , S を同定する。

$$\text{minimize } \|L\|_* + \lambda \|S\|_1 \quad \text{subject to } L + S = X \quad (5)$$

ここで、 λ は重みパラメータである。元のデータ行列 X の代わりに、低ランク行列 L に対して主成分分析を行い、次元削減を行うことで、データ行列 X の特徴量を、外れ値に依存しない、低次のモードとして抽出することができる。本研究では、実際には下記の特異値分解を適用し、次元縮約を行う。

$$L \approx \Phi_r \Sigma_r V_r^T = \Phi_r a \quad (6)$$

ここで Φ_r と V_r は、それぞれ左特異ベクトルと右特異ベクトルを列方向に並べた行列であり、 Σ_r は特異値が対角成分に大きいものから並んだ行列である。

3. 検証例題

1564 件の仮想的な南海トラフ地震・津波シナリオを作成し、観測点配置の最適設計のためにランダムに選定した 1414 件をデータ行列 X として用い、残る 150 件をテストシナリオとした津波予測¹⁾を行うことで観測点配置の妥当性の検証を行う。データ行列 X は、TUNAMI-N2³⁾ による津波シミュレーションから得られる $n = 134$ 点位置 (Fig. 2) における波高データとする。

データ行列に対して、PCA と RPCA によりそれぞれ求めた特異値の累積寄与率を Fig. 1 に示す。RPCA により算出した低次の空間モードは PCA の空間モードよりも高い寄与を有しており、データ行列全体に対して支配的な特徴量の抽出ができていくことがわかる。この結果から、式 (1) における次元削減モード数 r として、RPCA での特異値累積寄与率が 80 % となる $r = 7$ を用いることとした。 $r = 7$ とした空間モード行列 Φ_r を用い、最適化を行った結果得られた 7 点の配置を Fig. 2 に示す。PCA を用いた最適化では選定された観測点が沿岸部に集中しているが、RPCA に基づく配置ではそのような傾向が緩和されている。沖合での津波の伝播挙動を支配する低ランク成分とその特徴量が RPCA によって抽出され、合理的な最適化を実現している。

PCA と RPCA のそれぞれで得られた $p = 7$ の観測点を使用し、150 件の検証用シナリオに対する津波予測を実施した結果を Fig. 3, Fig. 4 に示す。10 分間の観測データからシナリオ同定された結果を基に、高知市沿岸部 (Point A) において予測される波高と実際に観測される波高 (検証用シナリオ) を示す。PCA による最適観測点の結果 (青線) と比較すると、RPCA を使用した場合の予測 (赤線) では 10 分間の観測情報から、約 6 時間分の波形をより高精度に予測していることがわかる。加えて、Point A における観測波高と予測波高の最大値を 150 件すべてでまとめた Fig. 4 からも、RPCA により最適化した観測網の方が全ての検証シナリオに対してより精度の高い予測をしていることがわかる。以上の結果から、RPCA により抽出された支配的な低次の空間モードと、それを参照した少数の最適観測点は、即時的な津波予測において有用であるといえる。

4. 結言

本研究では、ロバスト主成分分析 (RPCA) を使用した低ランク構造の抽出と、その特徴量に基づく最適化理論を津波波高データへと適用し、津波観測網の最適設計を行った。結果として、RPCA に基づく最適観測点配置は、PCA による沿岸部に集中した最適配置を緩和することが可能で

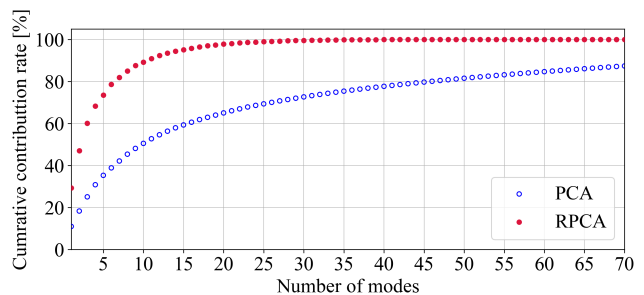


Fig. 1 Cumulative contribution ratio of singular values.

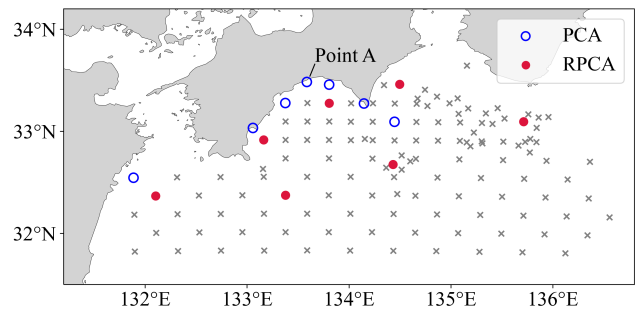


Fig. 2 Optimal configuration of $p = 7$ observation gauges

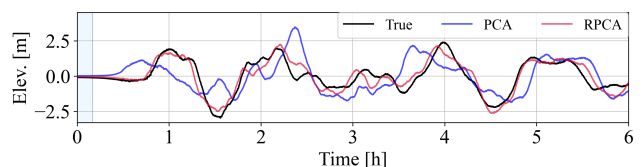


Fig. 3 Predicted and observed tsunami waveforms at Point A in Fig. 2 for one test scenario out of 150.

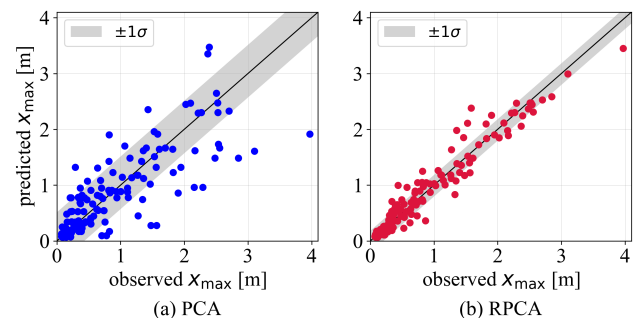


Fig. 4 Comparison of the predicted and observed maximum tsunami elevations at Point A in Fig. 2.

あることを確認した。また、最適観測網を用いた津波予測結果の比較から、RPCA を使用することにより、短時間の観測でより高精度な予測が可能であることを示した。

参考文献

- 1) R. Nomura, et al. : Sequential bayesian update to detect the most likely tsunami scenario using observational wave sequences, *Journal of Geophysical Research: Oceans*, Vol. 127, No. 10, p. e2021JC018324, 2022.
- 2) K. Manohar, et al. : Predicting shim gaps in aircraft assembly with machine learning and sparse sensing, *Journal of manufacturing systems*, Vol. 48, pp. 87–95, 2018.
- 3) F. Imamura: Review of tsunami simulation with a finite difference method, *Long-wave runup models*, pp.25–42, 1996.