Project Proposal: Bayesian Inference for Time Series and State-Space Models

1. Team Identification

- Weillin Sun, Yiwen Guo

https://github.com/ANNA-SUN-11/Stat447-team-project.git

2. Selected Project Theme

- Time Series and State-Space Bayesian Models
- Topic: how does government expenditure affect the Economy growth

3. Literature review summary

1) Wu, S.-Y., Tang, J.-H., & Lin, E. S. (2010). The impact of government expenditure on economic growth: How sensitive to the level of development? *Journal of Policy Modeling, 32*(6), 804-817. https://doi.org/10.1016/j.jpolmod.2010.05.011

- This paper finds the positive relationship between government expenditure and economic growth rate. But it only uses the linear regression model.

2) Almarashi, A. M., & Khan, K. (2020). Bayesian Structural Time Series. *Nanoscience and Nanotechnology Letters*, 12, 54–61. https://doi.org/10.1166/nnl.2020.3083

- This paper discussed Bayesian Structural Time Series (BSTS) models that provide more accurate long-term forecasting compared to classical ARIMA models when applied to stock price data.

3) West, M. (1995). Bayesian time series modeling and analysis. In *Proceedings of the XVth Workshop on Maximum Entropy and Bayesian Methods*. Santa Fe, New Mexico. Retrieved from http://www.isds.duke.edu

- This paper aimed to see if there are significant cyclic or quasi-periodic components in the deep-sea oxygen isotope time series data related to historical climate change by using DLMs and MCMC Methods.

For our topic, how does government expenditure affect the Economy growth? There are some papers that answer it in a linear regression model, but not many solve it by Time Series and State-Space Bayesian Models. We will try the model mentioned in the paper above.

4. Real-World Candidate Dataset

1) government expenditure measured by General government final consumption expenditure (current US$) | Data from World Bank Open Data and economic growth measured by GDP growth (annual %) | Data from World Bank Open Data. Then to make the data frame clear, we merged these two dataset together in R:

	Country.Name	year	government_expenditure	gdp_growth
1	Africa Eastern and Southern	X1961	1595460325	0.46873635
2	Australia	X1961	2227680356	2.48233047
3	Burundi	X1961	5250000	-13.74613505
4	Benin	X1961	23961099	3.14128046
5	Burkina Faso	X1961	26806827	4.04392798
6	Bangladesh	X1961	306058063	6.05816083
7	Bolivia	X1961	37500000	2.08929594
8	Brazil	X1961	2231397941	8.60000000
9	Botswana	X1961	6397423	6.34328370
10	Central African Republic	X1961	23648367	4.95355380
11	Canada	X1961	6591864065	3.16329168
12	Chile	X1961	509030476	5.24527188
13	China	X1961	7157317014	-27.27000000
14	Cote d'Ivoire	X1961	54847018	9.93255516
15	Congo, Rep.	X1961	20631155	8.35072672
16	Colombia	X1961	307899881	5.08921698
17	Costa Rica	X1961	56978340	1.87927656

2) Backup data: government expenditure measured by https://www.imf.org/external/datamapper/exp@FPP/USA/FRA/JPN/GBR/SWE/ESP/ITA/ZAF/IND?year=2023 from international monetary fund and economic growth measured by real GDP growth from https://www.imf.org/external/datamapper/NGDP_RPCH@WEO/OEMDC/ADVEC/WEOWORLD?year=2025. Then to make the data frame clear, we merged these two dataset together in R:

	country	year	government_expenditure	gdp_growth
1	Argentina	1980	15.173999786376999	0.69999999999999996
2	Australia	1980	33.210144042968999	2.8999999999999999
3	Austria	1980	50.008399963378999	2.2999999999999998
4	Belgium	1980	54.890598297118999	4.4000000000000004
5	Bolivia	1980	19.771600723266999	0.59999999999999998
6	Brazil	1980	6.8000001907348997	9.1999999999999993
7	Canada	1980	41.554401397705	2.2000000000000002
8	Chile	1980	23.840000152588001	7.9000000000000004
9	Colombia	1980	9.6000003814696999	4.4000000000000004
10	Costa Rica	1980	25.045099258423001	0.80000000000000004
11	Denmark	1980	52.661998748778998	-0.69999999999999996
12	Dominican Republic	1980	12.836400032043001	8
13	Eswatini	1980	15.305765567006	-3.7999999999999998
14	Ethiopia	1980	13.496405302758999	4

5. Approach and Methodology

Aimed to see if the Bayesian Structural Time Series (BSTS) models provide more accurate long-term forecasting compared to classical ARIMA models when applied to the economic growth rate.

1) **Data Preprocessing:**
   - Cleaning and structuring time series data for analysis.
   - Handling missing values and potential anomalies.
2) **ARIMA (AutoRegressive Integrated Moving Average)**
   - Determine the best ARIMA parameters (p, d, q) using Auto-ARIMA.
   - Fit the model and forecast the economic growth rate.
   - Evaluate the performance using Root Mean Squared Error (RMSE).
3) **Bayesian Structural Time Series (BSTS)**
   - Specify model components (trend, seasonality, regression terms).
   - Use Markov Chain Monte Carlo (MCMC) to estimate parameters.
   - Perform posterior inference and forecast.

Finally, Comparing the accuracy between ARIMA and BSTS model

6. Plan for teamwork

We will do the literature review, data preprocessing for 2 papers each person; then one person will do the ARIMA model for our topic and write the final paper, and the other one will do the BSTS model and comparison between two models. For the discussion and conclusion part, we will do it together.