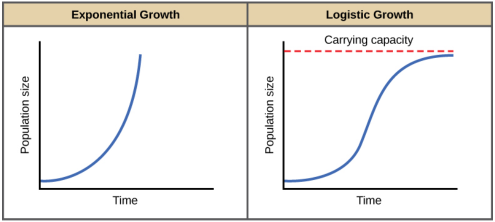
Sequential Monte Carlo Methods for United Kingdom Population Modelling

Kerry Liu

March 20, 2024

Introduction to Population Modeling

Population dynamics can be captured through models like the **Exponential Growth Model** ($P(t) = P_0 e^{rt}$) and the **Logistic Growth Model** [Newman et al., 2014], which introduces a carrying capacity to moderate growth. These models are foundational for understanding changes in population size influenced by birth, death, and immigration processes.



Challenges in Population Modeling

Objective: Accurately model true population dynamics $\{x_t\}$ using available data $\{y_t\}$ — census, births, deaths, immigration.

Challenge:

- Data is often incomplete or imperfect
- Simplistic models may not capture the complex reality of population dynamics, necessitating a more robust modeling approach.

Need for SSMs: The limitations of traditional models and data inaccuracies highlight the need for State Space Models (SSMs), offering a flexible framework to reconcile observed data with underlying population processes.

Introducing State Space Models (SSMs)

State Space Models (SSMs) [Auger-Méthé et al., 2021] employ a dual-equation framework to refine population modeling:

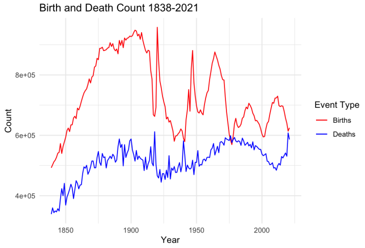
- **Process (State Transition) Equation:** $x_t = f(x_{t-1}, \eta_t)$ [1] models the evolution of the true population state (x_t), incorporating process noise (η_t) to account for randomness in demographic events.
- **Observation (Measurement) Equation:** $y_t = g(x_t, \epsilon_t)$ links the unobserved population states to observed data (y_t), accounting for observation errors (ϵ_t).

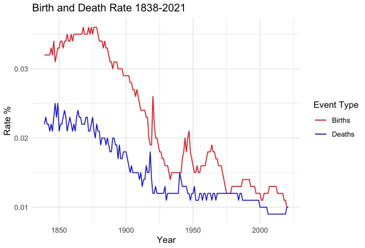
SSMs bridge the gap between theoretical models and real-world data. They allow for the specification of a variety of functions for $f()$ and $g()$, as well as the distributions of η_t and ϵ_t. This flexibility enables the fitting or forecasting of population dynamics with greater accuracy.

[1] Assume first-order Markov Chain here

Data Overview

- Population estimates for England & Wales (1838-2021), adjusted for census accuracy.
- Annual birth and death counts (1838-2021).
- Migration data by the Office for National Statistics (ONS), available from 1964 onwards.





Fitting SSMs: The Frequentist Approach

State Space Models (SSMs) can be fitted using the Frequentist inference, focusing on estimating the hidden states based on observed data

Key Concepts:

- **Joint Likelihood (L_j)**: The probability of observing the sequence of data ($y_{1:T}$) and the hidden states ($x_{1:T}$), defined as $L_j = \prod_{t=1}^{T} g(y_t|x_t, \theta_o) f(x_t|x_{t-1}, \theta_p)$.
- **Conditional Likelihood**: The conditional distribution of the states given the observations up to time t called filtering ('forward') or, given all T obs, as smoothing ('backwards')
- **Marginal Likelihood (L_m)**: The likelihood of the parameters given the observed data, integrating out the hidden states. It's crucial for model comparison and parameter estimation.

First estimate parameters θ by maximizing the (L_m), then use the MLE estimator $\hat{\theta}$ in the conditional likelihood to estimate states

Bayesian Inference for SSMs

The Bayesian approach treats both the states and parameters as random variables, incorporating prior knowledge through distributions.

Key Concepts:

- **Posterior Distribution**: Updated beliefs about the state and parameters after observing data, combining the likelihood function with prior distributions using Bayes' theorem and is commonly estimated using Monte Carlo Methods.

Challenge: Intractable Likelihood

- Frequentists: High-dimensional integrals of marginal likelihood
- Bayesian: Direct computation of posterior distributions $p(x_{1:t}, \theta | y_{1:t})$ with increasing data dimensions

Inference:

- The **Kalman** and **Particle filters** can be used to recursively update the estimate of the state given new observations.

Kalman Filter Equations and Model Specification

The Kalman Filter is a recursive (analytical) solution for the linear dynamic system estimation problem [Harvey, 1990]. It operates in two fundamental steps: prediction and update, allowing it to refine state estimates with each new observation.

Process Equation:

$$x_t = c_t + T_t \cdot x_{t-1} + \eta_t, \quad \eta_t \sim \mathcal{N}(0, H_t)$$

Observation Equation:

$$y_t = x_t + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, G_t)$$

Model Parameters:

- c_t: Annual net migration count.
- T_t: State transition factor (growth rate).
- η_t, ϵ_t: Process and observation noise, respectively.

Slide 6: Model Specification for Kalman Filter

The Kalman Filter updates estimates of the system's state using linear projections and the innovations from observed data:

For each step at time $t-1$, the Kalman filter:

1. Given state mean x_{t-1} and variance P_{t-1}
2. **Prediction**
   2.1 we predict the state as $\hat{x}_t = T_t x_{t-1} + c_t$
   2.2 with the variance as $\hat{P}_t = T_t P_{t-1} T_t^T + H_t$ where $\hat{P}_t$ is the predicted variance at time t.
3. **Update**
   3.1 We then calculate the Innovation (Prediction Error) as $e_t = y_t - \hat{x}_t = y_t - T_t \hat{x}_{t-1} - c_t$.
   3.2 The Kalman Gain (adjustment term) $K_t = \hat{P}_t + G_t$
   3.3 We then update our estimation of x_t with a normal distribution of mean $x_t = \hat{x}_t + \hat{P}_t {K_t}^{-1} e_t$ and variance: $P_t = (I - \hat{P}_t {K_t}^{-1})\hat{P}_t$

Optimizing Process Error (H_t) in the Kalman Filter

Objective: As studies suggest, incorrect estimation of parameters propagate to estimation of states [Newman et al., 2014]. We don't have the true process error variance (H_t), hence use **MLE** estimate

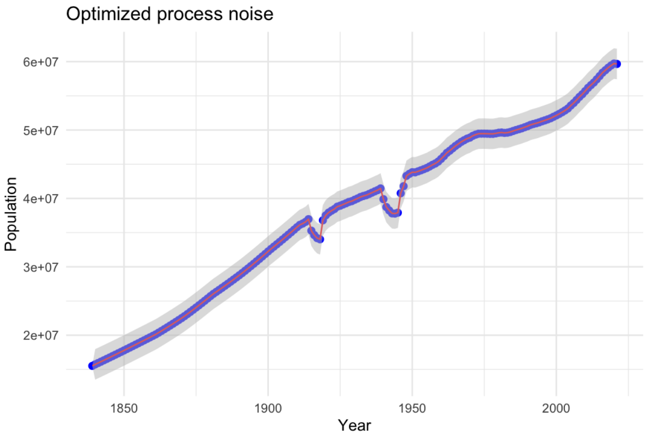
Methodology:

1. Define the *Process* and *Observation Equations* in Slide 7 to compute the joint likelihood (L_j).
2. To Maximize the marginal likelihood (L_m) of observed data w.r.t H_t, use **BFGS algorithm** to minimize the negative log-likelihood $-\mathcal{L}(\bar{H}_t, G_t, Y)$, estimating H_t without direct computation of the Hessian [Nocedal and Wright, 2006].

BFGS Steps:

1. Initialization: Set G_t as 3% of the mean population count and the initial guess of H_t with the variance of the yearly difference in population.
2. Iterative updates the estimation H_t through gradient information and the BFGS approximation of the inverse Hessian (2nd-order partial derivative) till convergence.

Kalman Filter Results



Transition matrix $T_t = 1.01$, Initial variance $P_0 = 10^3$, process error variance H_t from default 1.61×10^{11} to optimized 1.77×10^{11}, and Observation error variance $G_t = 1.13 \times 10^6$

Sequential Monte Carlo & Bootstrap Filter: Model Basis

Context: Employing a Bayesian state-space approach to capture non-linear, non-Gaussian population dynamics.

SBI Model Components:

- **Survival Process:** Modeled as Binomial deaths with constant death rate ϕ, influencing surviving population count.
- **Birth Process:** Modeled as Poisson births with rate $\beta \times x_t^1$ the surviving population, added to the population.
- **Immigration:** A fixed value input, directly increasing the population count.

SSMs Equations:

$$x_t \sim \text{Poisson}(x_{t-1} - \text{Binomial}(x_{t-1}, \phi) + \text{Poisson}(\beta \times x_t^1) + I_t)$$
$$y_t \sim \text{Lognormal}(\log(x_t), \log(G_t))$$

Poisson process equation allow for demographic stochasticity, and observations modeled as Lognormal to account for density-dependent variance [Hostetler and Chandler, 2015].

Bootstrap Filter Procedure

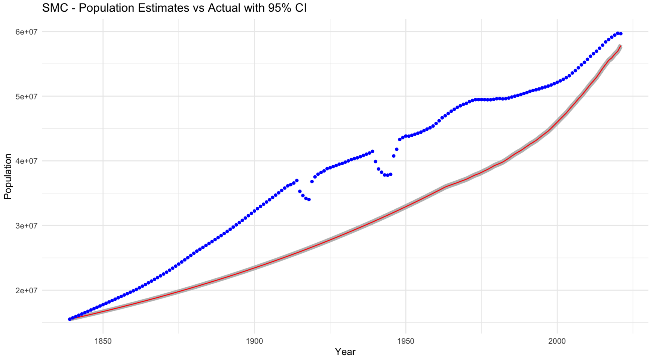
Sequential Monte Carlo Method: A "online" [2]approach to estimate posterior distributions in a Bayesian state-space model.
Procedure:

1. **Sample** new particles based on the state transition equation
2. **Weighting** particles according to how well they align with current observations, normalized sample weight to sum to one.
3. **Resampling** draws a new set of particles proportional to their weights, focusing on more likely states.
4. **Approximation** of the posterior distribution with the updated set of particles for current time step.

Implementation: Utilizing `NimbleSMC` for efficient computation within the NIMBLE framework, leveraging R for accessibility and C++ for performance.
Challenges: Particle degeneracy - few particles dominate the posterior estimate, require strategies to ensure particle diversity.

[2] iterative predict the current states without recompute whole data set

Bootstrap Filter Results



$\phi/\beta \sim \text{Uniform}(0, 0.1)$, with initial population $x_0 \sim \text{Poisson}(\lambda = x_0)$ and employs 10^4 particles for the Bootstrap Filter, with 95% posterior density intervals in grey

Particle Filter with Dynamic Birth and Death Rates

Objective: Enhance variability by incorporating dynamic recalculations of birth and death rates, as well as observation error variance at each time t into the bootstrap filter algorithm.

State Space Model:

$$x_t = x_{t-1} - \text{Binomial}(x_{t-1}, \phi_{t-1}) + \text{Poisson}(x_t^1 \cdot \beta_{t-1}) + I_t + \epsilon_{t,i},$$
$$y_t \sim \text{Lognormal}(\log(x_t), \log(G_{t,t})),$$

- x_t: Estimated population, with process error $\epsilon_{t,i} \sim \mathcal{N}(0, H_t)$.
- y_t: Observed population, with dynamically calculated observation error variance $G_{t,t}$ as 3% of previous year population count.

Dynamic Rates:

$$\beta_{t,i} \sim \text{Uniform}(0, 0.1),$$
$$\phi_{t,i} \sim \text{Uniform}(0, 0.1).$$

- Birth and death rates initialized from a uniform distribution, reflecting additional stochasticity.

Particle Filter with Dynamic Rates: Filtering Procedure

Simulation & Weighting:

1. Deaths and births simulated for each particle, updating population estimates.
2. Weights updated based on the log-normal distribution of observed versus estimated populations.

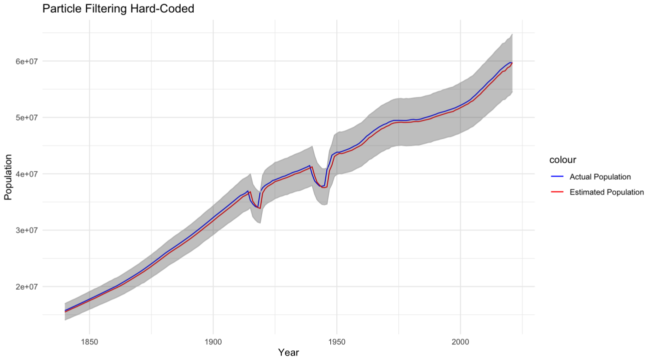
Resampling Strategy:

- If the Effective Sample Size (ESS) falls below 80%, particles are resampled.
- Resampling ensures diversity in particle set, preventing weight collapse.

Outcome:

- Posterior mean serves as point estimator; 95% posterior density intervals capture uncertainty.
- Dynamic rates allow the model to adapt more flexibly to new data, potentially improving accuracy over static-rate models.

Dynamic Bootstrap Filter Results



10^4 particles and an 80% ESS resampling threshold

Slide 12: Conclusion and Future Work

Conclusion:

- Our study demonstrates the robustness of State Space Models (SSMs) in modeling complex population dynamics, addressing limitations of traditional deterministic models.
- Through the application of Frequentist and Bayesian approaches, including Kalman and Particle filters, we've shown how SSMs can adapt to and reconcile imperfect and incomplete population data.
- The introduction of dynamic birth and death rates in our third model presents a significant improvement in capturing the stochastic nature of population processes, closely aligning model estimates with observed data.

Future Work:

- **Geographical Expansion:** Aim to analyze population dynamics separately for England, Scotland, Wales, and Northern Ireland to understand regional demographic trends better.
- **Data Integration:** Leverage comprehensive datasets from the National Records of Scotland and NISRA, incorporating detailed time series on births, deaths, migration, and population estimates since the mid-19th century.
- **Methodological Advancements:** Alongside Sequential Monte Carlo and Kalman filter methods, explore the use of Markov Chain Monte Carlo (MCMC) methods for model fitting. This approach aims to enhance our models' capacity to capture unique demographic trends and influences within each region.

References I

Marie Auger-Méthé, Ken Newman, Diana Cole, Fanny Empacher, Rowenna Gryba, Aaron A. King, Vianey Leos-Barajas, Joanna Mills Flemming, Anders Nielsen, Giovanni Petris, and Len Thomas. A guide to state–space modeling of ecological time series. *Ecological Monographs*, 91(4):e01470, 2021. doi: 10.1002/ecm.1470.

Andrew C. Harvey. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge, 1990.

Jeffrey A. Hostetler and Richard B. Chandler. Improved state-space models for inference about spatial and temporal variation in abundance from count data. *Ecology*, 96(6):1713–1723, 2015. doi: https://doi.org/10.1890/14-1487.1. URL `https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/14-1487.1`.

K. B. Newman, S. T. Buckland, B. J. T. Morgan, R. King, D. L. Borchers, and D. J. Cole. *Modelling Population Dynamics: Model Formulation, Fitting and Assessment Using State-Space Methods*. Methods in Statistical Ecology. Springer, New York, NY, 1 edition, 2014. ISBN 978-1-4939-0976-6. doi: 10.1007/978-1-4939-0977-3.

References II

Jorge Nocedal and Stephen J. Wright. *Quasi-Newton Methods*, pages 135–163. Springer New York, New York, NY, 2006. ISBN 978-0-387-40065-5. doi: 10.1007/978-0-387-40065-5_6. URL https://doi.org/10.1007/978-0-387-40065-5_6.