

Examen de
candidatura para
optar por el grado y
título de Maestría
Académica en
Estadística

LA SOBREPAREMETRIZACIÓN EN EL ARIMA:
UNA APLICACIÓN A DATOS COSTARRICENSES

César Gamboa Sanabria

Agenda

 Introducción

 Fundamentos teóricos

 Metodología

 Resultados esperados

 Cronograma de trabajo

Progreso

Introducción

Fundamentos teóricos

Metodología

Resultados esperados

Cronograma de trabajo

Antecedentes

- Uso de datos temporales.

- Investigación académica.
- Finanzas.
- Investigación de mercados.
- Demografía.

- Proyecciones.

- Métodos de análisis

- Los modelos $ARIMA(p, d, q)(P, D, Q)_s$
- Box-Jenkins.

El problema

- La identificación.
 - Visual.
- Alternativas.
 - `auto.arima()` (R. Hyndman & Khandakar, 2008)
 - Criterios de selección.
- Someter a prueba más modelos.

Objetivos

Objetivo general

Proponer un algoritmo alternativo más exhaustivo para la selección de modelos ARIMA mediante la sobreparametrización de los términos de la ecuación del ARIMA.

Objetivos específicos

1. Generar los escenarios de estimación de los distintos modelos ARIMA mediante permutaciones de los términos (p, d, q) y (P, D, Q) para la estimación de los posibles procesos que gobiernan una determinada serie temporal.
2. Aplicar diversos métodos de validación en la estimación de procesos que gobiernan la serie cronológica.
3. Contrastar la precisión de la estimación así como la generación de pronósticos con otros métodos similares, aplicados en datos costarricenses.
4. Integrar el desarrollo de la metodología de análisis de series temporales en una librería del lenguaje estadístico R.

Importancia y aportes del estudio

- ¿Campos que se benefician?
 - Demografía.
 - Actuarial.
 - Política pública.
- Comparación procesos que gobiernan la serie cronológica.

Progreso

Introducción

Fundamentos teóricos

Metodología

Resultados esperados

Cronograma de trabajo

Fundamentos teóricos

- Qué es una serie de tiempo (Hipel & McLeod (1994)).
- Dos grandes formas de estudio en las series de tiempo:
 - Componentes (Hernández (2011)).
 - Tendencia-ciclo, estacionalidad, irregularidad.
 - Autocorrelaciones.
- Series aditivas o multiplicativas.
- Supuestos (Agrawal & Adhikari (2013)).
- Identificación del modelo (Hyndman & Athanasopoulos, 2018).

Modelos ARIMA

- Modelos Autorregresivos (AR) (Box et al., 1994).
- Modelos de medias móviles (MA).
- Modelos Autorregresivos de Medias Móviles (ARMA).
- Modelos Autorregresivos Integrados de Medias Móviles (ARIMA).
- Funciones
 - Autocorrelación.
 - Autocorrelación parcial.
- Autocorrelogramas.
- La sobreparametrización y el análisis combinatorio.

Progreso

Introducción

Fundamentos teóricos

Metodología

Resultados esperados

Cronograma de trabajo

Materiales

- Series reales.
 - Tasa de mortalidad infantil interanual.
 - Mortalidad por causa externa.
 - Incentivos salariales del sector público.
 - Intereses y comisiones del sector público.
- Herramientas analíticas y de procesamiento.
- Series simuladas.
 - Valores aleatorios.
 - Selección del proceso.
 - Selección de los coeficientes.
 - Ajuste del modelo.
 - Simulación de la serie.

Métodos

- Análisis exploratorio.
- Partición de los datos.
- Estimación de modelos.
 - `auto.arima()`.
 - Sobreparametrización.
 - ARIMA estándar.
- Análisis de los errores.
- Pronósticos.
- Medidas de bondad de ajuste (Adhikari et al. (2013)).
 - AIC, AICc, BIC.
- Medidas de precisión.
 - MAE, MASE, RMSE.
- Tiempo de procesamiento.

Progreso

Introducción

Fundamentos teóricos

Metodología

Resultados esperados

Cronograma de trabajo

Resultados esperados: un espejo de la metodología aplicado a escenarios simulados y series reales

- Análisis exploratorio.
- Partición de los datos.
- Estimación de modelos.
 - `auto.arima()`.
 - Sobreparametrización.
 - ARIMA estándar.
- Análisis de los errores.
- Pronósticos.
- Medidas de bondad de ajuste.
 - AIC, AICc, BIC.
- Medidas de precisión.
 - MAE, MASE, RMSE.
- Tiempo de procesamiento.

Tiempos de estimación

Cuadro 1: Tiempos de estimación en minutos del modelo sugerido según el método de estimación para cada serie cronológica.

Método de estimación	Series simuladas						Series reales			
	ARIMA(1,0,0)	ARIMA(1,0,1)	ARIMA(2,0,3)	ARIMA(4,0,2)	ARIMA(0,0,1)(0,1,1)[12]	ARIMA(2,1,4)(3,0,3)[12]	TMII	EXTERNA	INCENTIVOS	INTERESES
auto.arima()	0.1056	0.0425	0.0783	0.1097	0.1625	5.0880	3.9190	0.1412	2.8172	0.6145
Sobreparametrización	8.2360	8.1266	5.2904	6.7233	39.5444	26.8193	53.1779	46.0911	21.4405	32.5420
ARIMA estándar	0.0064	0.0045	0.0044	0.0047	0.2976	0.1979	0.3798	0.2142	0.1049	0.1080

Fuente: Elaboración propia a partir de procesos de estimación.

Progreso

Introducción

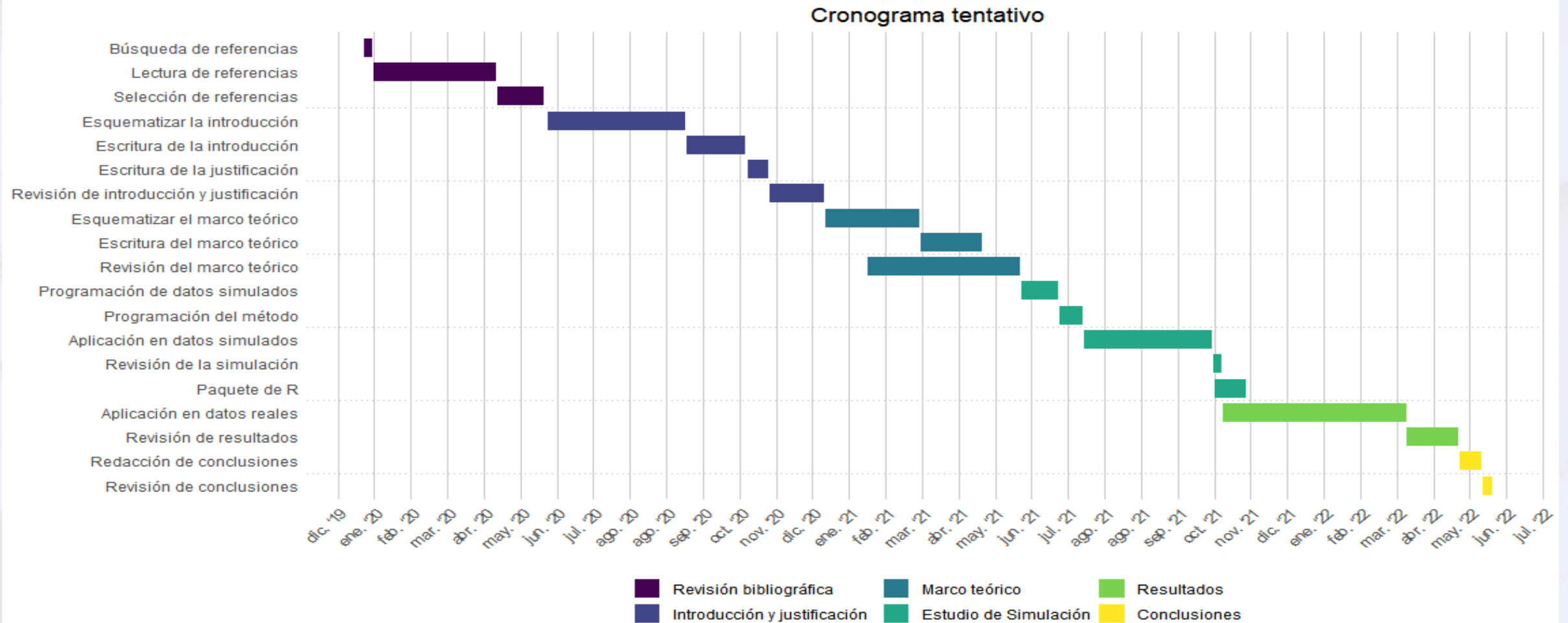
Fundamentos teóricos

Metodología

Resultados esperados

Cronograma de trabajo

Siguientes etapas



Referencias

- Hyndman, R., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for r. *Journal of Statistical Software, Articles*, 27 (3), 1–22. <https://doi.org/10.18637/jss.v027.i03>
- Rosero-Bixby, L. (2018). *Producto c para SUPEN. Proyección de la mortalidad de Costa Rica 2015-2150*. CCP-UCR. <http://srv-website.cloudapp.net/documents/10179/999061/Nota+t%C3%A9cnica+tablas+de+vida+segunda+parte>
- Hipel, K. W., & McLeod, A. I. (1994). *Time series modelling of water resources and environmental systems*. Elsevier Science. <https://books.google.co.cr/books?id=t1zG8OUbgdgC>
- Hernández, O. (2011). *Introducción a las series cronológicas* (1st ed.). Editorial Universidad de Costa Rica. <http://www.editorial.ucr.ac.cr/ciencias-naturales-y-exactas/item/1985-introduccion-a-las-series-cronologicas.html>
- Agrawal, R., & Adhikari, R. (2013). An introductory study on time series modeling and forecasting. *Nova York: CoRR*.
- Hyndman, R. J., & Athanasopoulos, G. (2018b). *Forecasting: Principles and practice*. OTexts. https://books.google.co.cr/books?id=_bBhDwAAQBAJ
- Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). *Time series analysis: Forecasting and control*. Prentice Hall. <https://books.google.co.cr/books?id=sRzvAAAAMAAJ>
- Adhikari, R., K, A. R., & Agrawal, R. K. (2013). *An introductory study on time series modeling and forecasting* (pp. 42–45). Lap Lambert Academic Publishing GmbH KG. <https://arxiv.org/ftp/arxiv/papers/1302/1302.6613.pdf>

Muchas gracias por su
atención.
