

Métodos en inferencia bayesiana causal

Dr. Matias Lopez-Rosenfeld (Profesor adjunto, DC, FCEN, UBA) con la colaboración del Dr. Gustavo Landried (Ayudante de Primera, DC, FCEN, UBA)

Programa:

Esta materia está enfocada en la evaluación de argumentos causales alternativos mediante (aproximaciones a) el sistema de razonamiento de las ciencias con datos: la aplicación estricta de las reglas de la probabilidad. El objetivo principal es revisar los métodos desarrollados en las últimas décadas para: especificar matemáticamente los argumentos causales expresados en lenguaje natural mediante métodos gráficos intuitivos; precisar cómo la estructura causal influye en el flujo de inferencia entre las variables del modelo; identificar el efecto causal entre variables de un modelo causal en base a datos observacionales; diseñar experimentos que permitan evaluar modelos causales alternativos; y tomar decisiones óptimas en contextos de incertidumbre.

Temario:

1. Modelos gráficos y evaluación de modelo causal en contextos constantes
 - Validación de proposiciones. Reglas de probabilidad: preservación de la creencia. Previa y predicción con todas las hipótesis. Especificación gráfica con redes bayesianas. Evaluación de modelos causales.
 - Taller: Generación de bases de datos sintéticas y evaluación de modelos causales alternativos.
 - Bibliografía sugerida: Samaja (cap 3.1-4), Klimovsky (cap 4), Bishop [6] (sec 1-4).
2. Aplicación estricta de las reglas de la probabilidad
 - Distribuciones conjugadas. Ejemplos: modelos de recomendación; Monty Hall extendido; modelo de habilidad. Estimaciones en línea en series temporales. Inferencia exacta en modelos de regresión lineal.
 - Taller: Implementación de modelos recomendación Beta-Binomial, Monty Hall Categórica-Dirichlet, habilidad-desempeño Gaussiana-Gaussiana. Estimación temporal en-línea y regresión lineal bayesiana.
 - Bibliografía sugerida: Bishop [2] (cap 2.1-3).
3. Emergencia del sobreajuste por selección y el balance natural por evaluación
 - Problemas computacionales de las reglas de probabilidad. Evaluación de modelos lineales y sobreajuste. Procesos Gaussianos.
 - Taller: Evaluación de modelos lineales y predicción con ensambles de modelos. Implementación de procesos gaussianos.
 - Bibliografía sugerida: Bishop [2] (sec 1.1-3, 3.3-4, 6.4.1-2).
4. Sorpresa: el problema de la comunicación con la realidad
 - Comunicación con la realidad, función de costo, entropía y sorpresa. Distribuciones de probabilidad máxima entropía.
 - Taller: Simulación numérica de apuestas, optimización de apuesta, minimización de divergencia KL.
 - Bibliografía sugerida: Klimovsky (cap 2), Samaja (sec 3.5-3.6), MacKay (sec 1.1, 2.4-6, 4.1), Kelly.

5. Especificación de modelo por Factor Graph e inferencia por pasaje de mensajes
 - Modelos causales probabilísticos con factor graph, sum-product algorithm, estimador ELO bayesiano.
 - Taller: Implementación de mensajes en modelo de habilidad y algoritmo suma-producto.
 - Bibliografía sugerida: Bishop (sec 8.2.1-2), Neal (cap 3).
6. Teorías causales
 - Modelos $A \rightarrow B$ y $B \rightarrow A$ en contextos dinámicos, teorías causales con gates, simulación de datos.
 - Taller: Generación de datos sintéticos y evaluación de teorías causales alternativas.
 - Bibliografía sugerida: Neal (sec 2.1, 4.1), Winn.
7. Niveles de razonamiento causal y estimación de efecto causal
 - Paradoja de Yule-Simpson, razonamiento causal, backdoor criterion, estimación del efecto causal.
 - Taller: Inferencia del efecto causal en modelos causales lineales.
 - Bibliografía sugerida: Neal (cap 4), Cinelli.
8. Semana disponible para prácticas, exámenes o compensación por retrasos o feriados.
9. Series temporales e inferencia en modelos de historia completa
 - Propagación de información forward-backward, loopy belief propagation, estimación de habilidad.
 - Taller: Implementación de algoritmo forward-backward, estudio de bases de datos de tenis, estimación de sesgo en Monty Hall temporal.
 - Bibliografía sugerida: Dangauthier, Bishop [2] (sec 13.2.2-4).
10. Inferencia causal en series temporales
 - Series de tiempo jerárquicas, teorías alternativas de aprendizaje, evaluaciones contrafactuales.
 - Taller: Implementación y evaluación de teorías de aprendizaje, evaluación del efecto causal de intervenciones.
 - Bibliografía sugerida: Brodersen, Bishop [2] (sec 13.3).
11. Aproximación de la inferencia por paseos al azar en el espacio de hipótesis
 - Generadores de números pseudo-aleatorios, Monte Carlo, Metropolis Hasting, nociones de Gibbs Sampling y Hamiltonian Monte Carlo. Lenguajes de programación probabilística.
 - Taller: Implementación de sampling, Monte Carlo, modelos gráficos en lenguaje probabilístico.
 - Bibliografía sugerida: Bishop [2] (sec 11.1-3), Stan user (ejemplos), Martin (ejemplos)
12. Aproximadores de eventos raros por algoritmos de Monte Carlo
 - Simulación de eventos raros, Sequential Monte Carlo, importance sampling y resampling, filtrado de partículas.
 - Taller: Experimentos en modelos State-Space, estimación de evidencia.
 - Bibliografía sugerida: Chopin (cap 10).

13. Evaluación de modelos mediante algoritmos de Monte Carlo
 - Evaluación numérica de modelos alternativos, predicción en modelos de historia completa, SMC combinado con MCMC, Ibis.
 - Taller: Evaluación de modelos causales temporales, estimación de evidencia.
 - Bibliografía sugerida: Chopin (cap 12, 17).
14. Isomorfismo probabilidad-evolución y apuestas óptimas
 - Procesos de selección probabilística y evolutiva, criterio Kelly, fractional kelly, diversificación, cooperación, especialización, otras.
 - Taller: Simulación de apuestas binario, evaluación de estrategias y criterios de apuesta.
 - Bibliografía sugerida: Czege, Peters, Kelly, Koller (cap 23).
15. Hackatón "apuestas de vida": un problema de acción-percepción
 - Ciclos de acción-percepción, problema similar al Monty Hall extendido, competencia de inferencia.
 - Taller: Hackatón.

Bibliografía Adicional:

1. Halpern JY. Reasoning about uncertainty. 2nd ed. MIT press; 2017.
2. Bishop CM. Pattern recognition and machine learning. Springer; 2006.
3. Winn J. Causality with gates. In: Artificial Intelligence and Statistics. PMLR; 2012. p. 1314–1322.
4. Samaja J. Epistemología y metodología: elementos para una teoría de la investigación científica. EUDEBA; 1999.
5. Klimovsky G. Las desventuras del conocimiento científico; 1994.
6. Bishop CM. Model-based machine learning. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2013;371(1984):20120222.
7. MacKay DJ. Information theory, inference and learning algorithms. Cambridge university press; 2003.
8. Kelly Jr JL. A New Interpretation of Information Rate. Bell System Technical Journal. 1956.
9. Neal B. Introduction to causal inference. Course Lecture Notes (draft). 2020.
10. Cinelli C, Forney A, Pearl J. A crash course in good and bad controls. Sociological Methods & Research. 2022; p. 00491241221099552.
11. Dangauthier P, Herbrich R, Minka T, Graepel T. Trueskill through time: Revisiting the history of chess. In: Advances in Neural Information Processing Systems; 2008. p. 337–344.
12. Brodersen KH, Gallusser F, Koehler J, Remy N, Scott SL. Inferring causal impact using Bayesian structural time-series models. The Annals of Applied Statistics. 2015.
13. Team SD. Stan User's Guide.
14. Martin OA, Kumar R, Lao J. Bayesian Modeling and Computation in Python. CRC Press; 2022.
15. Chopin N, Papaspiliopoulos O, et al. An introduction to sequential Monte Carlo. vol. 4. Springer; 2020.
16. Czege D, Giaffar H, Tenenbaum JB, Szathmáry E. Bayes and Darwin: How replicator populations implement Bayesian computations. BioEssays. 2022; p. 2100255.

17. Peters O. The ergodicity problem in economics. *Nature Physics*. 2019;15(12):1216–1221.
18. Koller D, Friedman N. *Probabilistic graphical models: principles and techniques*. MIT press; 2009.
19. Jaynes ET. *Bayesian methods: General background*; 1984.
20. McElreath R. *Statistical rethinking: A Bayesian course with examples in R and Stan*. CRC press; 2020.
21. Pearl J. *Causality*. Cambridge university press; 2009.1956.