

INSTITUTO TECNOLÓGICO AUTÓNOMO DE MÉXICO
EST-46114: Estadística Multivariada y Datos Categóricos (Maestría)
EST-24114: Análisis Multivariado (Licenciatura)
Inferencia Bayesiana en Alta Dimensión

Prof. Juan Carlos Martínez Ovando
`juan.martinez.ovando@itam.mx`

Enero-Junio 2016

Objetivo:

Este es un curso que contempla el estudio y aplicación de métodos estadísticos y de aprendizaje de maquina contemporáneos para analizar datos de gran escala (muchas dimensiones, muchas variables). Se estudian las nociones teóricas de los métodos de regularización para modelos sobre-parametrizados, prestando particular atención a su implementación bajo el enfoque bayesiano de inferencia. El curso contempla también el estudio de algunos modelos bayesianos no paramétricos también empleados en el análisis de grandes cantidades de datos.

Prerrequisitos Licenciatura: Estadística Aplicada II, Estadística Bayesiana (deseable, mas no indispensable).

Prerrequisitos Maestría: Fundamentos de Estadística, Modelos Lineales Generalizados (deseable) y Estadística Computacional (deseable).

Temas:

1. Introducción
 - a) ¿Por qué los modelos de alta dimensión?
 - b) ¿Por qué el enfoque bayesiano de inferencia?
2. Paradigma bayesiano de inferencia
 - a) Fundamentos y aprendizaje (Lindley, 1957; Gelman et al., 2004)
 - b) Herramientas computacionales (Robert and Casella, 2004)
 - c) Modelos jerárquicos de regresión (datos escalares) (Gelman and Hill, 2007)
 - d) Modelos jerárquicos de regresión (datos categóricos) (Gelman and Hill, 2007)
 - e) Comparación de modelos (Barbieri and Berger, 2004)
3. Modelos lineales de regresión
 - a) Modelos jerárquicos de regresión (datos escalares) (Gelman and Hill, 2007)
 - b) Modelos jerárquicos de regresión (datos categóricos/clasificación) (Gelman and Hill, 2007)
 - c) Comparación y diagnóstico ()
4. Regresión con muchas variables (*large p*, *small n*) y regularización
 - a) Escases y regularización (Tibshirani, 1996; Griffin and Brown, 2013)
 - b) Reducción de dimensiones (Componentes principales, factores latentes) (Press, 1982; West, 2003)

- c) Regresión penalizada (*ridge*, *shrinkage*, *lasso*) (Lengand et al., 2006; Park and Casella, 2008; Griffin and Brown, 2012)
 - d) Selección de variables (grupos de covariables) (George and McCulloch, 1997; Huang et al., 2012)
5. Aproximación de funciones no lineales
- a) Nociones de espacios funcionales (Härdle, 2004)
 - b) Regresión via *splines* y *smoothing splines* (DiMatteo et al., 2001)
 - c) Regresión via onduletas y bases radiales (Holmes and Mallick, 1998)
6. Predicción
- a) Comparación de modelos de regresión (Gutiérrez-Peña, 1997)
 - b) Selección estocástica de variables (Clyde, 1999; O'Hara and Sillanpää, 2009)
 - c) Promedio bayesiano de modelos (Clyde, 1999; Brown et al., 2002)
 - d) Bootstrap e inferencia bayesiana

Evaluación:

Un examen parcial (40 %) y un examen final (60 %).

Referencias

- Barbieri, M. M. and Berger, J. O. (2004). Optimal predictive model selection. *The Annals of Statistics*, 32(3):870–897.
- Brown, P. J., Vannucci, M., and Fearn, T. (2002). Bayes model averaging with selection of regressors. *Journal of the Royal Statistical Society, Series B*, 64(3):519–536.
- Clyde, M. A. (1999). Bayesian model averaging and model search strategies (with discussion). In Bernardo, J. M., Berger, J. O., Dawid, A. P., and Smith, A. F. M., editors, *Bayesian Statistics 6*, pages 157–186. Claredon Press, Oxford.
- DiMatteo, I., Genovese, C. R., and Kass, R. E. (2001). Bayesian curve-fitting with free-knot splines. *Biometrika*, 88(4):1055–1071.
- Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004). *Bayesian Data Analysis*. Chapman & Hall, London.
- Gelman, A. and Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, Cambridge.
- George, E. I. and McCulloch, R. E. (1997). Approaches for bayesian variable selection. *Statistica Sinica*, 7:339–373.
- Griffin, J. E. and Brown, P. J. (2012). Structuring shrinkage: Some correlated priors for regression. *Biometrika*, 99(2):481–487.
- Griffin, J. E. and Brown, P. J. (2013). Some priors for sparse regression modelling. *Bayesian Analysis*, 8(3):691–702.

- Gutiérrez-Peña, E. (1997). A Bayesian semiparametric approach to variable selection and model comparison in regression. In *Proceedings of the 51st Biennial Session of the International Statistical Institute, Invited Paper, Book 1*. International Statistical Institute.
- Härdle, W. (2004). *Applied Nonparametric Regression*. Bambridge University Press, Cambridge.
- Holmes, C. C. and Mallick, B. M. K. (1998). Bayesian radial basis functions of variable dimension. *Neural Computation*, 10(5):1217–1233.
- Huang, J., Breheny, P., and Ma, S. (2012). A selective review of group selection in high-dimensional models. *Statistical Science*, 27(4):481–499.
- Lengand, C., Lin, Y., and Wahba, G. (2006). A note on the LASSO and related procedures in model selection. *Statistica Sinica*, 16:1273–1284.
- Lindley, D. V. (1957). A statistical paradox. *Biometrika*, 44(1-2):187–192.
- O’Hara, R. B. and Sillanpää, M. J. (2009). A review of Bayesian variable selection methods: What, how and which. *Bayesian Analysis*, 4(1):85–118.
- Park, T. and Casella, G. (2008). The Bayesian lasso. *Journal of the American Statistical Association*, 103(482):681–686.
- Press, S. J. (1982). *Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference*. Krieger Publishing, Malabar, second edition.
- Robert, C. P. and Casella, G. (2004). *Monte Carlo Statistical Methods*. Springer, New York.
- Tibshirani, T. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*, 58(1):267–288.
- West, M. (2003). Bayesian factor regression models in the ”large p , small n ” paradigm. In Bernardo, J. M., Bayarri, M. J., Berger, J. O., Dawid, A. P., Heckerman, D., Smith, A. F. M., and West, M., editors, *Bayesian Statistics 7*, London. Oxford University Press.