emlogit: ECM Algorithm for Multinomial Logit

Soichiro Yamauchi*
Harvard University

March 14, 2020

Abstract

Abstract here

1 Introduction

The multinomial regression is a useful model for analyzing relationships between categorical/multinomial outcomes and a set of predictors. The multinomial logitstic regression (or multinomial logit) is one of the multinomial regression models often used in applied works. The model is also useful as a building block for more complicated Bayesian models. For example, the multinomial logit specification shows up anytime researchers want to incorporate covariates into their latent variable models (i.e., finite mixture models), such as stochastic blockmodels for network data or topic models for text data analysis.

In this note, I derive the Expectation and Conditional Maximization (ECM) algorithm for the standard multinomial logistic regression model as a useful reference and provide a **R** implementation as `emlogit` package. Since the seminal work by Polson et al on the Polya-Gamma representation of binomial family,

(also see Durante et al. 2018).

2 Model

*Graduate student, Department of Government, Harvard University. Email: syamauchi@g.harvard.edu