Contribution ID : **20**   Type : **not specified**

Regularized Joint Mixture Models

giovedì 29 giugno 2023 15:00 (20)

Regularized regression models are well studied and, under appropriate conditions, offer fast and statistically interpretable results. However, large data in many applications are heterogeneous in the sense of harboring distributional differences between latent groups. Then, the assumption that the conditional distribution of response Y given features X is the same for all samples may not hold. Furthermore, in scientific applications, the covariance structure of the features may contain important signals and its learning is also affected by latent group structure. We propose a class of mixture models for paired data (X, Y) that couples together the distribution of X (using sparse graphical models) and the conditional Y | X (using sparse regression models). The regression and graphical models are specific to the latent groups and model parameters are estimated jointly (hence the name "regularized joint mixtures"). This allows signals in either or both of the feature distribution and regression model to inform learning of latent structure and provides automatic control of confounding by such structure. Estimation is handled via an expectation-maximization algorithm, whose convergence is established theoretically. We illustrate the key ideas via empirical examples. An R package is available at https://github.com/k-perrakis/regjmix.

Keywords: distribution shifts, heterogeneous data, joint learning, latent groups, mixture models, sparse regression

Primary author(s) : PERRAKIS, Konstantinos (Department of Mathematical Sciences, Durham University); LARTIGUE, Thomas (Aramis Project Team, Inria & Center of Applied Mathematics, CNRS, Ecole Polytechnique, IP Paris); DONDELINGER, Frank (Lancaster Medical School); MUKHERJEE, Sach (German Center for Neurodegenerative Diseases, Bonn, Germany & MRC Biostatistics Unit, University of Cambridge)

Presenter(s) : PERRAKIS, Konstantinos (Department of Mathematical Sciences, Durham University)

Session Classification : Fifth Session