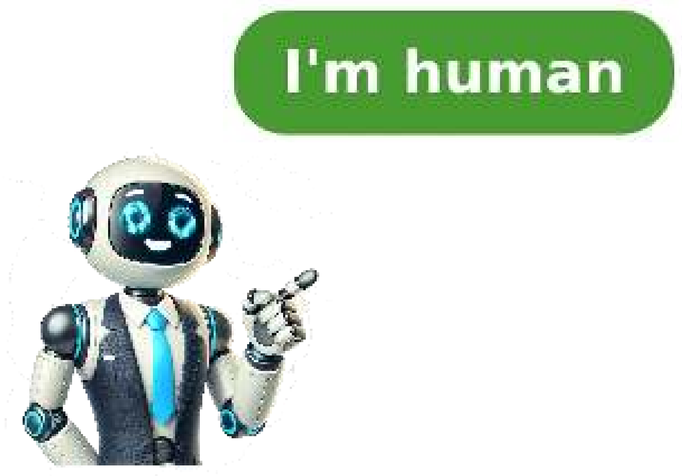
I'm human

Information geometry and alternating minimization procedures pdf

The Expectation-Maximization (EM) algorithm is widely used for statistical model parameter estimation, but its variants that don't exactly follow the EM formulation are also crucial in practice due to convergence issues and computational complexity. These variants deviate from the standard EM procedure by either approximating the E-step or employing incremental estimations. Since these modified procedures don't fit within the traditional EM framework, their convergence properties can't be directly inferred from the standard (G)EM results. This paper presents an information geometric approach to analyzing and understanding the behavior of such modified algorithms. Specifically, it applies this framework to two types of variants: incremental EM and variational EM. For incremental EM, conditions under which these algorithms converge in likelihood are discussed, while for variational EM, the role of E-step approximation in preventing convergence to local maxima in likelihood is highlighted. Given article text here In the field of machine learning and statistics, researchers have proposed various algorithms for efficient likelihood estimation and reconstruction in emission tomography. Notable contributions include the use of incremental EM (Expectation Maximization) algorithms, such as COSEM, which ensures fast and global convergence. Additionally, graphical models have been explored through variational methods, as presented by Jordan et al. in "Learning in Graphical Models." The efficiency of likelihood estimators has also been studied, with Lehmann's work highlighting the importance of convex statistical distances. Other notable contributions include the development of the EM algorithm and its extensions, as well as fast improvements to the EM algorithm, such as Meilijson's method. Furthermore, researchers have investigated the convergence properties of the EM algorithm and its variants, including bound optimization algorithms, acceleration methods for large databases, and robust grassmann manifold convex hull collaborative representation learning. Recent works have also focused on kernel extensions for image set analysis, context-patch representation learning with adaptive neighbor embedding for face image super-resolution, and shuffled linear regression with outliers in both covariates and responses. Karimi et al. (2019) explore the global convergence of incremental expectation maximization methods, which are widely used in machine learning and neural networks. The authors propose a new approach to improve the efficiency of these methods. Xie and Nie (2019) develop a novel algorithm for achieving proportionality in user equilibrium traffic assignment, an important problem in transportation science. Weinberger et al. (2019) introduce k-vectors, an alternating minimization algorithm for learning regression functions, which has applications in communication, control, and computing. These studies build upon the foundation of information geometry and Bayesian methods, as discussed by Amari and Csiszár (2000), Bregman (1967), Chentsov (1982), Dempster et al. (1977), Eguchi (1983), and Xu (1997). Xu L. byy harmony learning, independent state space, and generalized apt financial analyses in IEEE transactions on neural networks 2001 12 4 822–849 Xu L. best harmony, unified rpcl and automated model selection for unsupervised and supervised learning on gaussian mixtures three-layer nets and me-rbf-svm models in international journal of neural systems 2001 11 1 43–69 Xu L. byy harmony learning structural rpcl and topological self-organizing on mixture models in neural networks 2002 15 8–9 1125–1151 Pearl J. probabilistic reasoning in intelligent systems networks of plausible inference san mateo ca morgan kaufmann 1988 Ikeda S, Tanaka T, Amari S. information geometry of turbo and low-density parity-check codes in IEEE transactions on information theory 2004 50 6 1097–1114 Ikeda S, Tanaka T, Amari S. stochastic reasoning free energy and information geometry neural computation 2004 16 9 1779–1810 Csiszár I. information measures a critical survey in transactions of the 7th prague conference 1974 83–86 Csiszár I. axiomatic characterizations of information measures entropy 2008 10 3 261–273 Ali M S, Silvey S D. a general class of coefficients of divergence of one distribution from another journal of the royal statistical society series b 1966 28 1 131–142 Amari S. α-divergence is unique belonging to both f-divergence and bregman divergence classes IEEE transactions on information theory 2009 55 11 4925–4931 Cichocki A, Adunek R, Phan A H, Amari S. nonnegative matrix and tensor factorizations john wiley 2009 Havrda J, Charvát F. quantification method of classification process concept of structural α-entropy kybernetika 1967 3 30–35 Chernoff H. a measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations annals of mathematical statistics 1952 23 4 493–507 Matsuyama Y. the α-em algorithm surrogate likelihood maximization using α-logarithmic information measures IEEE transactions on information theory 2002 49 3 672–706 Amari S. integration of stochastic models by minimizing α-divergence neural computation 2007 19 10 2780–2796 Given text appears to be a list of academic articles related to neural computation, information geometry, and blind source separation, featuring authors such as Minami M, Eguchi S, Amari S, and Yuille A L. The citations span from 1978 to 2004, indicating a comprehensive review of the field over several decades.