Accelerating Any EM Algorithm without Sacrificing Simplicity and Stability

Ravi Varadhan[1], Christophe Roland[2], and Hormuzd Katki[3]

The EM algorithm is a ubiquitous computational approach for obtaining maximum likelihood estimates in incomplete data problems. Primary reasons for its popularity are simplicity and stability. However, the stability of EM is usually achieved at a high cost of slow, linear convergence. This limits the usefulness of EM in problems with complex statistical models, high-dimensionality, and large scale data. We have recently (Varadhan and Roland, 2004; Roland and Varadhan 2005) developed two new classes of iterative schemes, Steffensen-type methods for EM (STEM) and squared iterative methods (SQUAREM), to accelerate the convergence of EM. SQUAREM schemes, obtained by ``squaring" the underlying STEM methods, are faster and more efficient. The proposed methods are completely general as they can accelerate any linearly convergent fixed point iteration, and hence, any EM-type algorithm. SQUAREM schemes exhibit either fast-linear or superlinear convergence (3 to 30-fold in our examples). We will illustrate the superior convergence behavior of SQUAREM with a simple binary Poisson mixture example. We will also demonstrate the usefulness of SQUAREM schemes on an important problem in population genetics – the reconstruction of haplotypes from population genotype data.

The proposed schemes are extremely easy to implement, since they work solely with the EM updates. Auxiliary quantities such as complete or observed data log-likelihood, gradient, and hessian are not required. Most importantly, SQUAREM schemes achieve impressive gains in speed without sacrificing the stability of EM. Although the proposed schemes are generally non-monotone, we have developed simple globalization strategies that leverage the base EM iteration to provide enhanced stability. These combined attributes of simplicity, stability, speed, and generality, make SQUAREM methods highly attractive as an off-the-shelf accelerator for any EM algorithm.

We will discuss important computational issues involved in the implementation and use of SQUAREM techniques. An implementation of the new methods is available for general use as an R function.

---

[1] Assistant Professor, School of Medicine, Johns Hopkins University, rvaradhan@jhmi.edu

[2] Post-Doctoral Fellow, Laboratoire Paul Painleve, UFR de Mathematiques Pures et Appliquees-M3, Universite des Sciences et Technologies de Lille, 59655 Villeneuve d'Ascq cedex, France, christophe.roland@math.univ-lille1.fr

[3] Mathematical Statistician, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, katkih@mail.nih.gov