

RESUMO

O presente trabalho tem como objetivo a apresentação do algoritmo EM, que é uma ferramenta computacional utilizada para o cálculo do estimador de máxima verossimilhança (EMV) de forma iterativa, principalmente em problemas envolvendo dados incompletos. Para isso precisamos obter o conjunto dos dados completos, que é o conjunto dos dados observados aumentado com o conjunto dos dados faltantes e a partir daí obter a função log-verossimilhança associada aos dados completos.

Sabemos que este algoritmo seguramente converge para o EMV e tem como base a ideia de substituir uma difícil maximização por uma sequência de maximizações mais fáceis, envolvendo dois passos, o passo "E" (esperança) que calcula o valor esperado do logaritmo da verossimilhança completa; e o passo "M", que encontra seu máximo. Os passos são repetidos até se atingir a convergência.

Antes de exemplificarmos o cálculo do EMV via algoritmo EM, definiremos os modelos hierárquicos que possuem como vantagem modelar processos complicados através de uma sequência de modelos relativamente simples, colocados em uma hierarquia. Além disso, lidar com hierarquia não é mais difícil que lidar com distribuições condicionais ou marginais.

Como exemplo de cálculo do EMV via algoritmo EM, utilizaremos modelos de regressão, onde os erros assumem distribuição t-student, que além de serem simétricos, possuem caudas mais pesadas, sendo o processo de estimação mais robusto (no sentido de acomodar valores extremos). Apresentaremos também aplicações para essa distribuição.

Palavras-chave: Estimação de Máxima Verossimilhança, Algoritmo EM, Modelos Hierárquicos, Modelo t-Student.