EM algorithm — Maximum likelihood estimation of the non-standard t-distribution

Lucas Støjko Andersen

October 9, 2022

1 The non-standard t-distribution

The non-standard t-distribution has density

$$g(x \mid \mu, \sigma^2, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu\sigma^2}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{(x-\mu)^2}{\nu\sigma^2}\right)^{-\frac{\nu+1}{2}}. \tag{1}$$

Maximum likelihood estimates have no closed analytic solutions.

2 Expectation Maximization

Consider (X, W) with joint density

$$f(x, w \mid \mu, \sigma^2, \nu) = \frac{1}{\sqrt{\pi\nu\sigma^2}\,2^{(\nu+1)/2}\Gamma(\nu/2)} w^{\frac{\nu-1}{2}} e^{-\frac{w}{2}\left(1+\frac{(x-\mu)^2}{\nu\sigma^2}\right)}. \tag{2}$$

Then $W \mid X$ has density

$$h(w \mid x) \propto w^{\frac{\nu+1}{2}-1} e^{-\frac{w}{2}\left(1+\frac{(x-\mu)^2}{\nu\sigma^2}\right)} \tag{3}$$

and so $W \mid X \sim \Gamma(\alpha, \beta)$ where

$$\alpha = \frac{\nu+1}{2} \qquad \beta = \frac{1}{2}\left(1 + \frac{(X-\mu)^2}{\nu\sigma^2}\right). \tag{4}$$

Therefore

$$E(W \mid X) = \frac{\alpha}{\beta} = \frac{\nu+1}{1 + \frac{(X-\mu)^2}{\nu\sigma^2}} \tag{5}$$

and

$$E(\log W \mid X) = \psi(\alpha) - \log(\beta) = \psi\left(\frac{\nu+1}{2}\right) - \log\left(\frac{1}{2}\left(1 + \frac{(X-\mu)^2}{\nu\sigma^2}\right)\right) \quad (6)$$

where ψ is the digamma function

$$\psi(x) = \frac{d}{dx}\log\Gamma(x). \quad (7)$$

It can easily be shown that the marginal distribution of X has a non-standard t-distribution by multiplying the joint density by $\Gamma(\alpha)/\beta^{\alpha}$.

2.1 Expectation step

Define H as the negative log-likelihood of the joint density:

$$\begin{aligned} H(\theta) = & -\frac{1}{2}\log\pi - \frac{1}{2}\log\nu - \frac{1}{2}\log\sigma^2 - \frac{\nu+1}{2}\log 2 - \log\Gamma(\nu/2) \\ & + \frac{\nu-1}{2}\log w - \frac{w}{2} - \frac{w}{2}\frac{(x-\mu)^2}{\nu\sigma^2} \end{aligned}$$

where $\theta = (\mu, \sigma^2, \nu)$. Define $Q(\theta \mid \theta') = E(H(X,W)|X)$ we have

$$\begin{aligned} Q(\theta \mid \theta') = & -\frac{1}{2}\log\pi - \frac{1}{2}\log\nu - \frac{1}{2}\log\sigma^2 - \frac{\nu+1}{2}\log 2 - \log\Gamma(\nu/2) \\ & + \frac{\nu-1}{2}C_1(\theta', X) \\ & - \frac{C_2(\theta', X)}{2} - C_2(\theta', X)\frac{(X-\mu)^2}{2\nu\sigma^2} \end{aligned}$$

where

$$C_1(\theta, x) = \psi\left(\frac{\nu+1}{2}\right) - \log\left(\frac{1}{2} + \frac{(x-\mu)^2}{2\nu\sigma^2}\right) \quad C_2(\theta, x) = \frac{\nu+1}{1 + \frac{(x-\mu)^2}{\nu\sigma^2}}.$$

2.2 Maximization (minimization) step