EM algorithm A: The EM algorithm for the t-distribution

Michaela Lukacova

2024-10-08

The marginal distribution of X is the t-distribution

The joint density of $Y = (X, W)$ is given by

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

Let $C = \frac{1}{\sqrt{\pi\nu\sigma^2}\Gamma(\nu/2)}$

$$f(x) = \int f(x, w) dw = C \int \frac{1}{2^{(\nu+1)/2}} w^{(\nu-1)/2} \exp\left(-\frac{w}{2}\left(1 + \frac{(x-\mu)^2}{\nu\sigma^2}\right)\right) dw$$

Let $z = \frac{\nu+1}{2}$, implying that $z - 1 = \frac{\nu-1}{2}$.

$$f(x) = \int \frac{1}{2^z} w^{z-1} \exp\left(-\frac{w}{2}\left(1 + \frac{(x-\mu)^2}{\nu\sigma^2}\right)\right) dw$$

And let $t = \frac{w}{2}\left(1 + \frac{(x-\mu)^2}{\nu\sigma^2}\right)$, implying that $w = 2t\left(1 + \frac{(x-\mu)^2}{\nu\sigma^2}\right)^{-1}$ and $dw = 2\left(1 + \frac{(x-\mu)^2}{\nu\sigma^2}\right)^{-1} dt$. Note that the limits of the integral are still the same ($(0, \infty)$).

$$f(x) = C \int \frac{1}{2^z} w^{z-1} \exp(-t) \, dw$$

Substituting w and $dw = 2\left(1 + \frac{(x-\mu)^2}{\nu\sigma^2}\right)^{-1} dt$ we obtain

$$f(x) = C \int \frac{1}{2^z}\left(2\left(1 + \frac{(x-\mu)^2}{\nu\sigma^2}\right)\right)^{-z} t^{z-1} \exp(-t) \, dt = C\left(1 + \frac{(x-\mu)^2}{\nu\sigma^2}\right)^{-z} \int t^{z-1} \exp(-t) \, dt$$

Recognizing the gamma function we finally obtain

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

Which is recognized as the density of the t-distribution.

Maximize the complete data log-likelihood and implement this as a function

We find the derivatives of the log-likelihood with respect to μ and σ^2 and set them equal to zero. This results in the following estimators.

```r
sample <- generator(100)
mu_hat <- mle_mu(sample$x, sample$w)
sigma2_hat <- mle_sigma2(sample$x, sample$w, mu_hat, 5)
mu_hat; sigma2_hat
```

```
## [1] 0.03768418
```

```
## [1] 0.7599045
```

EM algorithm

Suppose we only observe the X's. We wish to use the EM algorithm to estimate the parameters μ and σ^2. We denote the pair of parameters by $\theta = (\mu, \sigma^2)$. The E-step of the algorithm amounts to computing the expectation

$$E_{\theta'} \left(\log\left(f\left(x, w|\theta\right)\right)|X = x\right) = -n\log\sigma + \int \left(\frac{n(\nu-1)}{2}\log(w) - \sum_{i=1}^{n} \frac{w}{2}\left(1 + \frac{(x_i-\mu)^2}{\nu\sigma^2}\right)\right) f_{w|X=x}(w)dw$$

Note that $f_{w|X=x'}(w) = \frac{f(x',w)}{f(x')} \propto f(x', w)$. Nu mangler du bare at genkende den tæthed som en gamma tæthed...