Research Notes

Jacques Nel

April 29, 2020

Chapter 1

One

1.1 Lambert W transform

Lambert's W function is the inverse of $z = u\exp(u)$, ie. the function that satisfies

$$W(z)\exp(W(z)) = z \tag{1.1}$$

Let U be a continous RV with cdf $F_U(u|\boldsymbol{\beta})$, pdf $f_U(u|\boldsymbol{\beta})$, and parameter vector $\theta = (\boldsymbol{\beta}, \delta)$, then

$$Z = U\exp\left(\frac{\delta}{2}U^2\right), \quad \delta \in \mathbb{R} \tag{1.2}$$

is a non-central, non-scaled heavy tail Lambert $W \times F_X$ RV with parameter vector $\theta = (\boldsymbol{\beta}, \delta)$, where δ is the tail parameter.

For a continuous location-scale family RV X $F_X(x|\boldsymbol{\beta})$ define a location-scale heavy-tailed Lambert $W \times F_X$ RV

$$Y = \left\{U\exp\left(\frac{\delta}{2}U^2\right)\right\}\sigma_x + \mu_x, \quad \delta \in \mathbb{R} \tag{1.3}$$

Let X $F_X(x|\boldsymbol{\beta})$ be continuous scale-family RV, with standard deviation σ_x, let $U = X/\sigma_x$, then

$$Y = X\exp\left(\frac{\delta}{2}U^2\right), \quad \delta \in \mathbb{R}, \tag{1.4}$$

is a heavy-tailed Lambert $W \times F_X$ RV with parameter $\theta = (\boldsymbol{\beta}, \delta)$.

Let $\tau := (\mu_x(\boldsymbol{\beta}), \sigma_x(\boldsymbol{\beta}))$ define transformation eq. (1.3).

1.1.1 Inverse transform: "Gaussianize" heavy-tailed data

The inverse transformation of eq. (1.3) is

$$W_\tau(Y) := W_\delta\left(\frac{Y-\mu_x}{\sigma_x}\right)\sigma_x + \mu_x = U\sigma_x + \mu_x = X \tag{1.5}$$

with

$$W_\delta(z,\delta) := \operatorname{sgn}(z)\left(\frac{W(\delta z^2)}{\delta}\right)^{1/2}. \tag{1.6}$$

$W_\delta(z)$ is bijective for $\forall \delta > 0$ and $\forall z \in \mathbb{R}$.

1.1.2 Maximum Likelihood Estimation (MLE)

Let

$$z = \frac{y-\mu_x}{\sigma_x}, u = W_\delta(z), \text{ and } x = W_\tau(y) = u\sigma_x + \mu_x.$$

The cdf and pdf of a location-scale heavy tail lambert $W \times F_X$ RV Y are given by

$$G_Y(y|\boldsymbol{\beta},\delta) = F_X(W_\delta(z)\sigma_x + \mu_x|\boldsymbol{\beta}), \tag{1.7}$$

and

$$g_Y(y|\boldsymbol{\beta},\delta) = f_X\left(W_\delta\left(\frac{y-\mu_x}{\sigma_x}\right)\sigma_x + \mu_x|\boldsymbol{\beta}\right)\cdot\frac{W_\delta\left(\frac{y-\mu_x}{\sigma_x}\right)}{\frac{y-\mu_x}{\sigma_x}\left(1+W\left(\delta\left(\frac{y-\mu_x}{\sigma_x}\right)^2\right)\right)} \tag{1.8}$$

For i.i.d. sampe $\boldsymbol{y}$ $g_Y(y|\boldsymbol{\beta},\delta)$ the log-likelihood function is

$$l(\theta,\boldsymbol{y}) = \sum_{i=1}^{N}\log g_Y(y_i|\boldsymbol{\beta},\delta). \tag{1.9}$$

This is an MLE problem, ie.,

$$\hat{\theta}_{\text{MLE}} = \left(\hat{\boldsymbol{\beta}},\hat{\delta}\right)_{\text{MLE}} = \underset{\boldsymbol{\beta},\delta}{\operatorname{argmax}}\; l\left(\boldsymbol{\beta},\delta\middle|\boldsymbol{y}\right) \tag{1.10}$$

Equation (1.9) can be decomposed as

$$l\left(\boldsymbol{\beta},\delta\middle|\boldsymbol{y}\right) = l(\boldsymbol{\beta}|\boldsymbol{x}_\tau) + \mathcal{R}(\tau|\boldsymbol{y}) \tag{1.11}$$

where

$$l(\boldsymbol{\beta}|\boldsymbol{x}_\tau) = \sum_{i=1}^{N}\log f_X\left(W_\delta\left(\frac{y_i-\mu_x}{\sigma_x}\right)\sigma_x + \mu_x\middle|\boldsymbol{\beta}\right) = \sum_{i=1}^{N}\log f_X\left(\boldsymbol{x}_\tau|\boldsymbol{\beta}\right) \tag{1.12}$$

is the log-likelihood of the back-transformed data $\boldsymbol{x}_\tau$, and

$$\mathcal{R}(\tau|\boldsymbol{y}) = \sum_{i=1}^{N} \log R\left(\mu_x, \sigma_x, \delta|y_i\right) \tag{1.13}$$

with

$$R(\mu_x, \sigma_x, \delta|y_i) = \frac{W_\delta\left(\frac{y_i-\mu_x}{\sigma_x}\right)}{\frac{y-\mu_x}{\sigma_x}\left(1+\delta\left(W_\delta\left(\frac{y_i-\mu_x}{\sigma_x}\right)\right)^2\right)} \tag{1.14}$$

1.1.3 Gradient descent

Let $z_i = \frac{y_i-\mu_x}{\sigma_x}$.

$$\frac{\partial}{\partial\mu_x} l(\boldsymbol{\beta}|\boldsymbol{x}_\tau) = \sum_{i=1}^{N}\left[\left(1 - \frac{\sigma_x W(z_i^2)\operatorname{sgn}(y_i-\mu_x)}{\delta(y_i-\mu_x)\left(\frac{W(\delta z_i^2)}{\delta}\right)^{1/2}\left(1+W(\delta z_i^2)\right)\operatorname{sgn}\sigma_x} - \frac{\sigma_x\left(\frac{W(\delta z_i^2)}{\delta}\right)^{1/2} 2\gamma(y_i-\mu_x)}{\operatorname{sgn}\sigma_x}\right)\right.$$
$$\left.\Bigg/\left(\mu_x + \frac{\sigma_x\left(\frac{W(\delta z_i^2)}{\delta}\right)^{1/2}\operatorname{sgn}(y_i-\mu_x)}{\operatorname{sgn}\sigma_x}\right)\right]$$

$$\frac{\partial}{\partial\sigma_x} l(\boldsymbol{\beta}|\boldsymbol{x}_\tau) = \sum_{i=1}^{N}\left[\left(\frac{\left(\frac{W(\delta z_i^2)}{\delta}\right)^{1/2}\operatorname{sgn}(y_i-\mu_x)}{\operatorname{sgn}\sigma_x} - \frac{W(\delta z_i^2)\operatorname{sgn}(y_i-\mu_x)}{\delta\left(\frac{W(\delta z_i^2)}{\delta}\right)^{1/2}\left(1+W(\delta z_i^2)\right)\operatorname{sgn}\sigma_x}\right.\right.$$
$$\left.\left.-\frac{\sigma_x\left(\frac{W(\delta z_i^2)}{\delta}\right)^{1/2}\operatorname{sgn}(y_i-\mu_x)2\gamma(\sigma_x)}{\operatorname{sgn}^2\sigma_x}\right)\Bigg/\left(\mu_x + \frac{\sigma_x\left(\frac{W(\delta z_i^2)}{\delta}\right)^{1/2}\operatorname{sgn}(y_i-\mu_x)}{\operatorname{sgn}\sigma_x}\right)\right]$$

$$\frac{\partial}{\partial\delta} l(\boldsymbol{\beta}|\boldsymbol{x}_\tau) = \sum_{i=1}^{N}\left[\frac{\sigma_x\left(-\frac{W(\delta z_i^2)}{\delta^2} + \frac{W(\delta z_i^2)}{\delta^2(1+W(\delta z_i^2))}\right)}{2\left(\frac{W(\delta z_i^2)}{\delta}\right)^{1/2}\left(\mu_x + \frac{\sigma\left(\frac{W(\delta z_i^2)}{\delta}\right)^{1/2}\operatorname{sgn}(y_i-\mu_x)}{\operatorname{sgn}\sigma_x}\right)\operatorname{sgn}\sigma_x}\right]$$