Deming regression

MethComp package

May 2007

Anders Christian Jensen Steno Diabetes Center, Gentofte, Denmark
acjs@steno.dk

Contents

1 Introduction

This document is related to the `Deming` function in the package `MethComp` and contains the derivation of the maximum likelihood estimates related to the Deming regression model. It is based on the book 'Models in regression and related topics' (chapter three), from 1969 by Peter Sprent, but with more detailed calculations included.

2 Deming regression

The mathematical model $\eta = \alpha + \beta\xi$ describes a linear relationship between two variables ξ and η. Observations x and y of two variables are usually desribed by a regression of y on x where x is assumed to be observed without error (or, equivantly using the conditional distribution of y given x). In linear regression with observations subject to additive random variation on both x and y and observed values for individuals $(x_i, y_i), i = 1, \ldots, n$, a model may be written

$$x_i = \xi_i + e_{xi},$$

$$y_i = \eta_i + e_{yi} = \alpha + \beta\xi_i + e_{yi},$$

where e_{xi} and e_{y_i} denotes the random part of the model. This is known as a functional relationship because the ξ_i's are assumed to be fixed parameters, as oppposed to a structural relationship where some distribution for the ξ_i's is assumed. In the following it is assumed that the e_{xi}s are iid with $e_{yi} \sim N(0, \sigma^2)$, and that the e_{yi}s are iid with $e_{yi} \sim N(0, \lambda\sigma^2)$, for some $\lambda > 0$. Furthermore e_{xi} is assumed to be independent of e_{yi}.

The aim of this document is to derive the maximum likelihood estimates for α, β, ξ_i and σ^2 in the functional model stated above.

3 The likelihood function

The likelihood function $f_{x_1,x_2,\ldots,x_n,y_1,y_2,\ldots,y_n}(\alpha, \beta, \xi_1, \xi_2, \ldots, \xi_n, \sigma^2)$ denoted f is

$$f = \prod_{i=1}^{n} \left(2\pi\sigma^2\right)^{-\frac{1}{2}} \exp\left(-\frac{(x_i - \xi_i)^2}{2\sigma^2}\right) \left(2\pi\lambda\sigma^2\right)^{-\frac{1}{2}} \exp\left(-\frac{(y_i - \alpha - \beta\xi_i)^2}{2\lambda\sigma^2}\right)$$

and the loglikelihood, denoted L, is

$$\begin{aligned} L &= \sum_{i=1}^{n} -\frac{1}{2}\log\left(2\pi\sigma^2\right) - \frac{(x_i - \xi_i)^2}{2\sigma^2} - \frac{1}{2}\log\left(2\pi\lambda\sigma^2\right) - \frac{(y_i - \alpha - \beta\xi_i)^2}{2\lambda\sigma^2} \\ &= -\frac{n}{2}\log\left(4\pi^2\right) - \frac{n}{2}\log\left(\lambda\sigma^4\right) - \frac{\sum_{i=1}^{n}(x_i - \xi_i)^2}{2\sigma^2} - \frac{\sum_{i=1}^{n}(y_i - \alpha - \beta\xi_i)^2}{2\lambda\sigma^2}. \end{aligned}$$

It follows that the likelihood function is not bounded from above when σ^2 goes to 0, so in the following it is assumed that $\sigma^2 > 0$.

4 Solving for ξ_i

Differentiation of L with respect to ξ_i gives

$$\begin{aligned}\frac{\partial L}{\partial \xi_i} &= \frac{\partial}{\partial \xi_i}\left(-\frac{\sum_{i=1}^{n}(x_i-\xi_i)^2}{2\sigma^2} - \frac{\sum_{i=1}^{n}(y_i-\alpha-\beta\xi_i)^2}{2\lambda\sigma^2}\right) \\ &= \frac{(x_i-\xi_i)}{\sigma^2} + \frac{\beta(y_i-\alpha-\beta\xi_i)}{\lambda\sigma^2}.\end{aligned}$$

Setting $\frac{\partial L}{\partial \xi_i}$ equal to zero yields

$$\frac{\partial L}{\partial \xi_i} = 0 \;\;\Rightarrow\;\; \xi_i = \frac{\lambda\sigma^2 x_i + \beta\sigma^2 y_i - \beta\alpha\sigma^2}{\lambda\sigma^2 + \beta^2\sigma^2} = \frac{\lambda x_i + \beta(y_i-\alpha)}{\lambda+\beta^2}. \tag{1}$$

So to estimate ξ_i, estimates for β and α are needed. Therefore focus is turned to the derivation of $\hat{\alpha}$.

5 Solving for α

Differentiation of L with respect to α gives

$$\begin{aligned}\frac{\partial L}{\partial \alpha} &= \frac{\partial}{\partial \alpha}\left(-\frac{\sum_{i=1}^{n}(y_i-\alpha-\beta\xi_i)^2}{2\lambda\sigma^2}\right) \\ &= \frac{\sum_{i=1}^{n}(y_i-\alpha-\beta\xi_i)}{\lambda\sigma^2},\end{aligned}$$

and putting $\frac{\partial L}{\partial \alpha}$ equal to zero yields

$$\frac{\partial L}{\partial \alpha} = 0 \;\;\Rightarrow\;\; \alpha = \frac{1}{n}\sum_{i=1}^{n}(y_i - \beta\xi_i).$$

Now one can use (1) to dispense with ξ_i

$$\begin{aligned}
\alpha &= \frac{1}{n}\sum_{i=1}^{n}(y_i - \beta\xi_i) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \beta\frac{\lambda x_i + \beta(y_i - \alpha)}{\lambda + \beta^2}\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \beta\frac{\lambda x_i + \beta y_i}{\lambda + \beta^2} + \frac{\beta^2\alpha}{\lambda + \beta^2}\right) \\
&\Updownarrow \\
\alpha\left(1 - \frac{\beta^2}{\lambda + \beta^2}\right) &= \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \beta\frac{\lambda x_i + \beta y_i}{\lambda + \beta^2}\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(y_i\left(1 - \frac{\beta^2}{\lambda + \beta^2}\right) - x_i\frac{\beta\lambda}{\lambda + \beta^2}\right) \\
&\Updownarrow \\
\alpha &= \frac{1}{n}\sum_{i=1}^{n}\left(y_i - x_i\frac{\beta\lambda}{\lambda + \beta^2}\frac{\lambda + \beta^2}{\lambda}\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i\beta) \\
&= \overline{y} - \overline{x}\beta.
\end{aligned}$$

Hence the estimate for α becomes

$$\hat{\alpha} = \overline{y} - \overline{x}\hat{\beta}.$$

6 Solving for β

Differentiation of L with respect to β gives

$$\frac{\partial L}{\partial \beta} = \frac{\partial}{\partial \beta}\left(-\frac{\sum_{i=1}^{n}(y_i - \alpha - \beta\xi_i)^2}{2\lambda\sigma^2}\right) = \frac{\sum_{i=1}^{n}(y_i - \alpha - \beta\xi_i)\xi_i}{\lambda\sigma^2}.$$

Setting $\frac{\partial L}{\partial \beta}$ equal to zero yields

$$\frac{\partial L}{\partial \beta} = 0 \Leftrightarrow \sum_{i=1}^{n}(y_i - \alpha - \beta\xi_i)\xi_i = 0,$$

and using (1)

$$\begin{aligned}
0 &= \sum_{i=1}^{n}(y_i - \alpha - \beta\xi_i)\xi_i \\
&= \sum_{i=1}^{n}\left(y_i - \alpha - \beta\frac{\lambda x_i + \beta(y_i-\alpha)}{\lambda+\beta^2}\right)\frac{\lambda x_i + \beta(y_i-\alpha)}{\lambda+\beta^2}.
\end{aligned}$$

This implies that

$$\begin{aligned}
0 &= \sum_{i=1}^{n}\Big((y_i-\alpha)(\lambda+\beta^2) - \beta\lambda x_i - \beta^2(y_i-\alpha)\Big)\Big(\lambda x_i + \beta(y_i-\alpha)\Big) \\
&= \sum_{i=1}^{n}\lambda^2 x_i(y_i-\alpha) + \beta^2\lambda x_i(y_i-\alpha) - \beta\lambda^2 x_i^2 - \beta^2\lambda x_i(y_i-\alpha) + \\
&\quad \sum_{i=1}^{n}\beta\lambda(y_i-\alpha)^2 + \beta^3\lambda(y_i-\alpha)^2 - \beta^2\lambda x_i(y_i-\alpha) - \beta^3(y_i-\alpha)^2 \\
&= -\beta^2\lambda\Big(\sum_{i=1}^{n}x_iy_i\Big) - \beta\lambda^2\Big(\sum_{i=1}^{n}x_i^2\Big) + \lambda^2\Big(\sum_{i=1}^{n}x_iy_i\Big) \\
&\quad +\beta^2\lambda\alpha\Big(\sum_{i=1}^{n}x_i\Big) + \beta\lambda\Big(\sum_{i=1}^{n}(y_i-\alpha)^2\Big) - \lambda^2\alpha\Big(\sum_{i=1}^{n}x_i\Big).
\end{aligned}$$

Dividing with λ and using the fact that $\alpha = \overline{y}. - \overline{x}.\beta$ it is seen that

$$\begin{aligned}
0 &= -\beta^2\Big(\sum_{i=1}^{n}x_iy_i\Big) - \beta\lambda\Big(\sum_{i=1}^{n}x_i^2\Big) + \lambda\Big(\sum_{i=1}^{n}x_iy_i\Big) + \beta^2(\overline{y}. - \overline{x}.\beta)\Big(\sum_{i=1}^{n}x_i\Big) \\
&\quad +\beta\Big(\sum_{i=1}^{n}\big(y_i - (\overline{y}. - \overline{x}.\beta)\big)^2\Big) - \lambda(\overline{y}. - \overline{x}.\beta)\Big(\sum_{i=1}^{n}x_i\Big) \\
&= -\beta^2\Big(\sum_{i=1}^{n}x_iy_i\Big) - \beta\lambda\Big(\sum_{i=1}^{n}x_i^2\Big) + \lambda\Big(\sum_{i=1}^{n}x_iy_i\Big) + \beta^2\overline{y}.\Big(\sum_{i=1}^{n}x_i\Big) \\
&\quad -\beta^3\overline{x}.\beta\Big(\sum_{i=1}^{n}x_i\Big) + \beta\Big(\sum_{i=1}^{n}y_i^2\Big) + \beta\Big(\sum_{i=1}^{n}(\overline{y}. - \overline{x}.\beta)^2\Big) \\
&\quad -2\beta\Big(\sum_{i=1}^{n}y_i(\overline{y}. - \overline{x}.\beta)\Big) - \lambda\overline{y}.\Big(\sum_{i=1}^{n}x_i\Big) + \lambda\overline{x}.\beta\Big(\sum_{i=1}^{n}x_i\Big).
\end{aligned}$$

Splitting up the sums even more gives

$$\begin{aligned} 0 &= -\beta^2\Big(\sum_{i=1}^{n} x_i y_i\Big) - \beta\lambda\Big(\sum_{i=1}^{n} x_i^2\Big) + \lambda\Big(\sum_{i=1}^{n} x_i y_i\Big) + \beta^2\overline{y}.\Big(\sum_{i=1}^{n} x_i\Big) - \beta^3\overline{x}.\beta\Big(\sum_{i=1}^{n} x_i\Big) \\ &\quad +\beta\Big(\sum_{i=1}^{n} y_i^2\Big) + \beta\Big(\sum_{i=1}^{n} \overline{y}.^2\Big) + \beta\Big(\sum_{i=1}^{n} (\overline{x}.\beta)^2\Big) - 2\beta\Big(\sum_{i=1}^{n} \overline{y}.\overline{x}.\beta\Big) - 2\beta\Big(\sum_{i=1}^{n} y_i\overline{y}.\Big) \\ &\quad +2\beta\Big(\sum_{i=1}^{n} y_i\overline{x}.\beta\Big) - \lambda\overline{y}.\Big(\sum_{i=1}^{n} x_i\Big) + \lambda\overline{x}.\beta\Big(\sum_{i=1}^{n} x_i\Big). \end{aligned}$$

Finally the terms are sorted and collected according to powers of β:

$$\begin{aligned} 0 &= \beta^3\left(\sum_{i=1}^{n} \overline{x}.^2 - \overline{x}.\sum_{i=1}^{n} x_i\right) \\ &\quad +\beta^2\left(\overline{y}.\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i y_i - 2\sum_{i=1}^{n} \overline{y}.\overline{x}. + 2\sum_{i=1}^{n} y_i\overline{x}.\right) \\ &\quad +\beta\left(\sum_{i=1}^{n} y_i^2 - \lambda\sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} \overline{y}.^2 - 2\sum_{i=1}^{n} y_i\overline{y}. + \lambda\overline{x}.\sum_{i=1}^{n} x_i\right) \\ &\quad +\lambda\left(\sum_{i=1}^{n} x_i y_i - \overline{y}.\sum_{i=1}^{n} x_i\right). \end{aligned}$$

Since

- $\sum_{i=1}^{n} \overline{x}.^2 - \overline{x}.\sum_{i=1}^{n} x_i = 0$
- $\overline{y}.\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i y_i - 2\sum_{i=1}^{n} \overline{y}.\overline{x}. + 2\sum_{i=1}^{n} y_i\overline{x}. = -\mathrm{SPD}_{xy}$
- $\sum_{i=1}^{n} y_i^2 - \lambda\sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} \overline{y}.^2 - 2\sum_{i=1}^{n} y_i\overline{y}. + \lambda\overline{x}.\sum_{i=1}^{n} x_i = \mathrm{SSD}_y - \lambda\mathrm{SSD}_x$
- $\sum_{i=1}^{n} x_i y_i - \overline{y}.\sum_{i=1}^{n} x_i = \mathrm{SPD}_{xy}$

it is clear that the derivation of β comes down to solve

$$-\beta^2\mathrm{SPD}_{xy} + \beta(\mathrm{SSD}_y - \lambda\mathrm{SSD}_x) + \lambda\mathrm{SPD}_{xy} = 0. \tag{2}$$

For $\mathrm{SPD}_{xy} \neq 0$ this implies that

$$\begin{aligned} \beta &= \frac{-(\mathrm{SSD}_y - \lambda\mathrm{SSD}_x) \pm \sqrt{(\mathrm{SSD}_y - \lambda\mathrm{SSD}_x)^2 - 4(-\mathrm{SPD}_{xy})\lambda\mathrm{SPD}_{xy}}}{-2\mathrm{SPD}_{xy}} \\ &= \frac{\mathrm{SSD}_y - \lambda\mathrm{SSD}_x \pm \sqrt{(\mathrm{SSD}_y - \lambda\mathrm{SSD}_x)^2 + 4\lambda\mathrm{SPD}_{xy}^2}}{2\mathrm{SPD}_{xy}}. \end{aligned}$$

Since $\mathrm{SSD}_y - \lambda\mathrm{SSD}_x \leq \sqrt{(\mathrm{SSD}_y - \lambda\mathrm{SSD}_x)^2 + 4\lambda\mathrm{SPD}_{xy}^2}$ there is always a positive and a negative solution to (2). The desired solution should always have the same sign as SPD_{xy}, hence the solution with the positive numerator is selected. Therefore

$$\hat{\beta} = \frac{\mathrm{SSD}_y - \lambda\mathrm{SSD}_x + \sqrt{(\mathrm{SSD}_y - \lambda\mathrm{SSD}_x)^2 + 4\lambda\mathrm{SPD}_{xy}^2}}{2\mathrm{SPD}_{xy}}.$$

7 Solving for ξ_i - again

With estimates for β and α it is now possible to estimate ξ_i using (1):

$$\hat{\xi}_i = \frac{\lambda x_i + \hat{\beta}(y_i - \hat{\alpha})}{\lambda + \hat{\beta}^2}.$$

8 Solving for σ^2

Differentiation of L with respect to σ^2 gives

$$\begin{aligned}
\frac{\partial L}{\partial \sigma^2} &= \frac{\partial}{\partial \sigma^2}\left(-\frac{n}{2}\log(\lambda\sigma^4) - \frac{\sum_{i=1}^n (x_i - xi_i)^2}{2\sigma^2} - \frac{\sum_{i=1}^n (y_i - \alpha - \beta\xi_i)^2}{2\lambda\sigma^2}\right) \\
&= \frac{-n\sigma^2}{\sigma^4} + \frac{\sum_{i=1}^n (x_i - xi_i)^2}{2\sigma^4} + \frac{\sum_{i=1}^n (y_i - \alpha - \beta\xi_i)^2}{2\lambda\sigma^4} \\
&= \frac{-2\lambda n\sigma^2 + \lambda\sum_{i=1}^n (x_i - xi_i)^2 + \sum_{i=1}^n (y_i - \alpha - \beta\xi_i)^2}{2\lambda\sigma^4},
\end{aligned}$$

and setting $\frac{\partial L}{\partial \sigma^2}$ equal to zero yields

$$\begin{aligned}
\frac{\partial L}{\partial \sigma^2} = 0 \quad &\Rightarrow \quad -2\lambda n\sigma^2 + \lambda\sum_{i=1}^n (x_i - xi_i)^2 + \sum_{i=1}^n (y_i - \alpha - \beta\xi_i)^2 = 0 \\
&\Rightarrow \quad \sigma^2 = \frac{\lambda\sum_{i=1}^n (x_i - \xi_i)^2 + \sum_{i=1}^n (y_i - \alpha - \beta\xi_i)^2}{2\lambda n}.
\end{aligned}$$

To get a central estimate of σ^2 one must divide by $n-2$ instead of $2n$ since there are $n+2$ parameters to be estimated, namely $\xi_1, \xi_2, \ldots, \xi_n, \alpha$ and β. Hence the degrees of freedom are $2n - (n+2) = n - 2$. Therefore

$$\hat{\sigma}^2 = \frac{\lambda\sum_{i=1}^n (x_i - \hat{\xi}_i)^2 + \sum_{i=1}^n (y_i - \hat{\alpha} - \hat{\beta}\hat{\xi}_i)^2}{2\lambda(n-2)}.$$

9 Summing up

$$\begin{aligned}
\hat{\alpha} &= \overline{y} - \overline{x}\hat{\beta} \\
\hat{\beta} &= \frac{\mathrm{SSD}_y - \lambda\mathrm{SSD}_x + \sqrt{(\mathrm{SSD}_y - \lambda\mathrm{SSD}_x)^2 + 4\lambda\mathrm{SPD}_{xy}^2}}{2\mathrm{SPD}_{xy}} \\
\hat{\sigma} &= \sqrt{\frac{\lambda\sum_{i=1}^n (x_i - \hat{\xi}_i)^2 + \sum_{i=1}^n (y_i - \hat{\alpha} - \hat{\beta}\hat{\xi}_i)^2}{2\lambda(n-2)}} \\
\hat{\xi}_i &= \frac{\lambda x_i + \hat{\beta}(y_i - \hat{\alpha})}{\lambda + \hat{\beta}^2}
\end{aligned}$$

These formula are implemented in the `Deming` function in the `MethComp` package.

10 The Deming function

```
Deming <-
function( x, y, vr=sdr^2, sdr=sqrt(vr), boot=FALSE, alpha=0.05 )
{
if( missing( vr ) & missing( sdr ) ) var.ratio <- 1
else var.ratio <- vr
vn <- c( deparse( substitute( x ) ),
         deparse( substitute( y ) ) )
alfa <- alpha
dfr <- data.frame( x=x, y=y )
dfr <- dfr[complete.cases(dfr),]
x <- dfr$x
y <- dfr$y
n <- nrow( dfr )
SSDy <- var( y )*(n-1)
SSDx <- var( x )*(n-1)
SPDxy <- cov( x, y )*(n-1)
beta <- ( SSDy - var.ratio*SSDx +
          sqrt( ( SSDy - var.ratio*SSDx )^2 +
                4*var.ratio*SPDxy^2 ) ) / ( 2*SPDxy)
alpha <- mean( y ) - mean( x ) * beta
ksi <- ( var.ratio*x + beta*(y-alpha) )/(var.ratio+beta^2)
sigma.x <- ( var.ratio*sum( (x-ksi)^2 ) +
                       sum( (y-alpha-beta*ksi)^2 ) ) /
# The ML-estiamtes requires 2*n at this point bu we do not think we have that
# many observations so we stick to (n-2). Any corroboation from litterature?
           ( (n-2)*var.ratio )
sigma.y <- var.ratio*sigma.x
sigma.x <- sqrt( sigma.x )
sigma.y <- sqrt( sigma.y )
if( !boot ){
res <- c( alpha, beta, sigma.x, sigma.y )
names( res ) <- c( "alpha", "beta", "sigma.x", "sigma.y" )
res
}
else{
if( is.numeric( boot ) ) N <- boot else N <- 1000
res <- matrix( NA, N, 4 )
for( i in 1:N )
   {
   wh <- sample( 1:n, n, replace=TRUE )
   res[i,] <- Deming( x[wh], y[wh], vr=var.ratio, boot=FALSE )
   }
ests <- cbind( c(alpha,beta,sigma.x, sigma.y),
               se <- sqrt( diag( cov( res ) ) ),
               t( apply( res, 2, quantile, probs=c(0.5,alfa/2,1-alfa/2 ) ) ) )
rownames( ests ) <- c( "Intercept", "Slope", paste( "sigma", vn, sep="." ) )
colnames( ests )<- c("Estimate", "S.e.(boot)", colnames(ests)[3:5] )
ests
}
}
```