Reproduce example 9.2

EM algotithm

Let θ = $(\mu_1, \mu_2, ln\sigma_{11}, ln\sigma_{22}, Z_\rho)$. σ_{11}, σ_{22}, Z_ρ are the variance of y_1, y_2 and Fisher Z transformation of correlation coefficient ρ.

Preparation

Since thers is no missing value in y_1, we have:

$$\mu_1 = n^{-1} \sum_{i=1}^{n} y_{i1}$$

$$ln\sigma_{11} = ln\left(n^{-1} \sum_{i=1}^{n} y_{i1}^2 - \mu_1^2\right)$$

From Example 7.1 we have:

$$\beta_{21.1}^{(t)} = \sigma_{12}^{(t)} / \sigma_{11}$$

$$\beta_{20.1}^{(t)} = \mu_2^{(t)} - \frac{\sigma_{12}^{(t)}}{\sigma_{11}} \mu_1$$

$$\sigma_{22.1}^{(t)} = \sigma_{22}^{(t)} - \frac{\sigma_{12}^{(t)2}}{\sigma_{11}}$$

where $\sigma_{12}^{(t)} = \rho * \sqrt{\sigma_{11} * \sigma_{22}^{(t)}}$

E-step

$$E\left[\sum_{i=1}^{n} y_{i2} | Y_{obs}, \theta^{(t)}\right] = \sum_{i=1}^{r} y_{i2} + \sum_{i=r+1}^{n} \left(\beta_{20.1}^{(t)} + \beta_{21.1}^{(t)} y_{i1}\right)$$

$$E\left[\sum_{i=1}^{n} y_{i2}^2 | Y_{obs}, \theta^{(t)}\right] = \sum_{i=1}^{r} y_{i2}^2 + \sum_{i=r+1}^{n} \left[(\beta_{20.1}^{(t)} + \beta_{21.1}^{(t)} y_{i1})^2 + \sigma_{22.1}^{(t)}\right]$$

$$E\left[\sum_{i=1}^{n} y_{i1} y_{i2} | Y_{obs}, \theta^{(t)}\right] = \sum_{i=1}^{r} y_{i1} y_{i2} + \sum_{i=r+1}^{n} y_{i1} (\beta_{20.1}^{(t)} + \beta_{21.1}^{(t)} y_{i1})$$

M-step (synchronously)

Update $\theta^{(t+1)}$ with $\theta^{(t)}$ and the expectations above.

$$\mu_2^{(t+1)} = n^{-1} E\left[\sum_{i=1}^{n} y_{i2} | Y_{obs}, \theta^{(t)}\right]$$

$$ln\sigma_{22}^{(t+1)} = ln\left\{E\left[\sum_{i=1}^{n} y_{i2}^2 | Y_{obs}, \theta^{(t)}\right] - 2\mu_2^{(t)} E\left[\sum_{i=1}^{n} y_{i2} | Y_{obs}, \theta^{(t)}\right] + n\mu_2^{2(t)}\right\} - ln(n)$$

$$Z_\rho^{(t+1)} = \frac{1}{2} ln\left(\frac{1+\rho^{(t+1)}}{1-\rho^{(t+1)}}\right)$$

where $\rho^{(t+1)} = \frac{E\left[\sum_{i=1}^{n} y_{i1} y_{i2} | Y_{obs}, \theta^{(t)}\right] - \mu_2^{(t)} \sum_{i=1}^{n} y_{i1} - \mu_1 E\left[\sum_{i=1}^{n} y_{i2} | Y_{obs}, \theta^{(t)}\right] + n\mu_1\mu_2^{(t)}}{n\sqrt{\sigma_{11}\sigma_{22}^{(t)}}}$

M-step (asynchronously)

Update $\theta^{(t+1)}$ with $\theta^{(t)}$, the calculated $\theta^{(t+1)}$ and the expectations above.

$$\mu_2^{(t+1)} = n^{-1} E\left[\sum_{i=1}^{n} y_{i2} | Y_{obs}, \theta^{(t)}\right]$$

$$ln\sigma_{22}^{(t+1)} = ln\left\{E\left[\sum_{i=1}^{n} y_{i2}^2 | Y_{obs}, \theta^{(t)}\right] - 2\mu_2^{(t+1)} E\left[\sum_{i=1}^{n} y_{i2} | Y_{obs}, \theta^{(t)}\right] + n\mu_2^{2(t+1)}\right\} - ln(n)$$

$$Z_\rho^{(t+1)} = \frac{1}{2} ln\left(\frac{1+\rho^{(t+1)}}{1-\rho^{(t+1)}}\right)$$

where $\rho^{(t+1)} = \frac{E\left[\sum_{i=1}^{n} y_{i1} y_{i2} | Y_{obs}, \theta^{(t)}\right] - \mu_2^{(t+1)} \sum_{i=1}^{n} y_{i1} - \mu_1 E\left[\sum_{i=1}^{n} y_{i2} | Y_{obs}, \theta^{(t)}\right] + n\mu_1\mu_2^{(t+1)}}{n\sqrt{\sigma_{11}\sigma_{22}^{(t+1)}}}$

SEM algorithm

I tried to update $\theta^{(t+1)}$ synchronously. However, the matrix DM is different from the results in *Meng and Rubin(1991)*. Therefore, I simulated both synchronous and asynchronous algorithm and found out that they used the asynchronous M-step in EM algorithm.
The major problem in SEM is to define the stability of $DM^{(t)}$. I consider DM[t] stable if K DM$^{(t)}$ in a row have tiny difference (e.g. $||DM^{(i+1)} - DM^{(i)}||_\infty$<0.0001 ,∀ i=t,t+1,...t+K-1).

SEM results

My results compared to the results in *Meng and Rubin(1991)*:

$$DM^*_{M\&R} = \begin{bmatrix} 0.33333 & 0.05037 & -0.02814 \\ 1.44444 & 0.29894 & 0.01921 \\ -0.64222 & 0.01529 & 0.32479 \end{bmatrix} \Delta V^*_{M\&R} = \begin{bmatrix} 1.0858 & 0.1671 & -0.0933 \\ 0.1671 & 0.0286 & -0.0098 \\ -0.0933 & -0.0098 & 0.0194 \end{bmatrix}$$

$$DM^*_{mine} = \begin{bmatrix} 0.33334 & 0.05037 & -0.02815 \\ 1.44443 & 0.29893 & 0.01918 \\ -0.64221 & 0.01530 & 0.32480 \end{bmatrix} \Delta V^*_{mine} = \begin{bmatrix} 1.0858 & 0.1671 & -0.0934 \\ 0.1670 & 0.0285 & -0.0098 \\ -0.0934 & -0.0098 & 0.0194 \end{bmatrix}$$

My script almost produces the same DM and ΔV.

Parameter	μ_2	$ln\sigma 22$	Z_ρ
(my) ML estimate	49.33	4.74	-1.45
(M&R's) ML estimate	49.33	4.74	-1.45
(my) s.e. from SEM	2.73	0.37	0.274
(M&R's) s.e. from SEM	2.73	0.37	0.274

And it produce the exact same ML-estimates and asympotic standard error as Table 9.1.

Bootstrap

I bootsrapped the data for 500 times and applied EM algorithm on every bootstrap samples. It turns out that the estimates are close to the estimates from SEM. But the standard error estiamte from bootstrap is lower than the one from SEM.
Additionally, an annoying fact in the process is that small sample size often leads to bad convergence in EM algorithm. I had to enlarge the sampling size to $4n$ to avoid bad convergence. Luckliy, the outcome is fair and sensible.

Bootstrap results

Parameter	μ_2	$ln\sigma 22$	Z_ρ
estimate from Bootstrap	49.33	4.70	-1.46
(M&R's) ML estimate	49.33	4.74	-1.45
s.e. from Bootstrap	1.36	0.13	0.14
(M&R's) s.e. from SEM	2.73	0.37	0.274