# Covariance pattern in LMMstar

Brice Ozenne

October 14, 2022

# 1 LV pattern

## 1.1 Theory

Consider 4 timepoints. The traditional parametrisation of the residual variance-covariance matrix of a factor model is

$$\Omega = \begin{bmatrix} \omega_1^2 + \tau & . & . & . \\ \lambda_2\tau & \omega_2^2 + \lambda_2^2\tau & . & . \\ \lambda_3\tau & \lambda_2\lambda_3\tau & \omega_3^2 + \lambda_3^2\tau & . \\ \lambda_4\tau & \lambda_2\lambda_4\tau & \lambda_3\lambda_4\tau & \omega_4^2 + \lambda_4^2\tau \end{bmatrix}$$

LMMstar uses a different parametrisation with distinct parameters for the variance and the correlation:

$$\Omega = \begin{bmatrix} \sigma_1^2 & . & . & . \\ \rho_1\rho_2\sigma_1\sigma_2 & \sigma_2^2 & . & . \\ \rho_1\rho_3\sigma_1\sigma_3 & \rho_2\rho_3\sigma_2\sigma_3 & \sigma_3^2 & . \\ \rho_1\rho_4\sigma_1\sigma_4 & \rho_2\rho_4\sigma_2\sigma_4 & \rho_3\rho_4\sigma_3\sigma_4 & \sigma_4^2 \end{bmatrix}$$

The two parametrisation are equivalent when assuming the same sign for all correlation (e.g. all positive correlation).

$\boldsymbol{\omega}, \boldsymbol{\tau}, \boldsymbol{\lambda} \rightarrow \boldsymbol{\rho}, \boldsymbol{\sigma}$: the $\sigma$ values can be deduce from the diagonal of $\Omega$.
To get the $\rho$ value, we can first multiply $\rho_1\rho_2\sigma_1\sigma_2 = \lambda_2\tau$ by $\rho_1\rho_3\sigma_1\sigma_3 = \lambda_3\tau$ and divide by $\rho_2\rho_3\sigma_2\sigma_3 = \lambda_2\lambda_3\tau$:

$$\frac{\rho_1^2\rho_2\rho_3\sigma_1^2\sigma_2\sigma_3}{\rho_2\rho_3\sigma_2\sigma_3} = \frac{\lambda_2\lambda_3\tau^2}{\lambda_2\lambda_3\tau}$$

$$\rho_1^2\sigma_1^2 = \tau$$

$$\rho_1^2 = \frac{\tau}{\omega_1^2 + \tau}$$

More generally, denoting $\lambda_1 = 1$, from:

$$\sigma_i^2 = \omega_i^2 + \lambda_i^2 \tau$$
$$\rho_i \rho_j \sigma_i \sigma_j = \lambda_i \lambda_j \tau$$

we can deduce

$$\rho_i^2 \rho_j \rho_k \sigma_i^2 \sigma_j \sigma_k = \lambda_i^2 \lambda_j \lambda_k \tau^2$$
$$\rho_i^2 \sigma_i^2 = \lambda_i^2 \tau$$
$$\rho_i = \sqrt{\frac{\lambda_i^2 \tau}{\omega_i^2 + \lambda_i^2 \tau}}$$

Technical $\rho_i$ could be negative but here use the assumption of same (positive) sign for all correlations.

**$\boldsymbol{\rho}, \boldsymbol{\sigma} \rightarrow \boldsymbol{\omega}, \boldsymbol{\tau}, \boldsymbol{\lambda}$:** we can re-use the previous result:

$$\rho_i^2 = \frac{\lambda_i^2 \tau}{\omega_i^2 + \lambda_i^2 \tau} = \frac{\lambda_i^2 \tau}{\sigma_i^2}$$

So for $i = 1$:

$$\tau = \rho_1^2 \sigma_1^2$$

and otherwise:

$$\lambda_i^2 = \frac{\rho_i^2 \sigma_i^2}{\rho_1^2 \sigma_1^2}$$
$$\lambda_i = sign(\rho_1 \rho_i) \frac{\rho_i \sigma_i}{\rho_1 \sigma_1}$$

One can then deduce $\omega_i$:

$$\omega_i = \sqrt{\sigma_i^2 - \lambda_i \tau} = \sigma_i \sqrt{1 - \rho_i^2}$$

## 1.2 Example

Simulate data

```
library(lava)
library(LMMstar)

mSim <- lvm(c(Y1,Y2,Y3,Y4)~eta+age)
latent(mSim) <- ~eta

set.seed(10)
n <- 100
n.time <- length(endogenous(mSim))
dfW.sim <- cbind(id = paste0("Id",1:n), sim(mSim, n = n, latent = FALSE))
dfW.sim$id <- factor(dfW.sim$id, unique(dfW.sim$id))
head(dfW.sim)
```

```
   id         Y1          Y2         Y3         Y4        age
1 Id1  0.8087642  0.02821369  2.0055318 2.29256267  0.8694750
2 Id2  0.3174894  0.92111736  0.8326184 1.09215142 -0.6800096
3 Id3  0.9880281  1.31941524  3.7496337 1.72867315  0.1732145
4 Id4 -0.3524308  0.95831086  1.1187839 1.03908643 -0.1594380
5 Id5  0.3496855 -0.57807269 -1.0256767 0.18052490  0.7934994
6 Id6  0.1276581  0.30103845  0.2336854 0.06061876  1.6943505
```

Convert to long format:

```
dfL.sim <- reshape(dfW.sim, direction = "long", idvar = c("id","age"),
   varying = paste0("Y",1:4), sep="")
dfL.sim$time <- as.factor(dfL.sim$time)
rownames(dfL.sim) <- NULL
head(dfL.sim)
```

```
   id        age time          Y
1 Id1  0.8694750    1  0.8087642
2 Id2 -0.6800096    1  0.3174894
3 Id3  0.1732145    1  0.9880281
4 Id4 -0.1594380    1 -0.3524308
5 Id5  0.7934994    1  0.3496855
6 Id6  1.6943505    1  0.1276581
```

Fit LVM:

```
m.lvm <- lvm(c(Y1,Y2,Y3,Y4)~eta+age, eta ~ 0)
latent(m.lvm) <- ~eta
e.lvm <- estimate(m.lvm, data = dfW.sim)
logLik(e.lvm)
```

```
'log Lik.' -651.0478 (df=16)
```

Export coefficient by type:

```
mu.lvm <- coef(e.lvm)[c(paste0("Y",1:n.time),paste0("Y",1:n.time,"~age"))]
lambda.lvm <- c(1,coef(e.lvm)[paste0("Y",2:n.time,"~eta")])
tau.lvm <- coef(e.lvm)["eta~~eta"]
omega.lvm <- coef(e.lvm)[paste0("Y",1:n.time,"~~Y",1:n.time)]
list(mu = mu.lvm,
     lambda = lambda.lvm,
     tau = tau.lvm,
     omega = omega.lvm)
```

```
$mu
        Y1         Y2         Y3         Y4     Y1~age     Y2~age     Y3~age 
-0.1835368 -0.1491306 -0.0194078  0.1459640  0.9502774  1.0535363  0.9671297 
    Y4~age 
 1.0349377 

$lambda
             Y2~eta    Y3~eta    Y4~eta 
1.0000000 0.8653753 1.1024519 1.0537868 

$tau
eta~~eta 
1.259395 

$omega
   Y1~~Y1    Y2~~Y2    Y3~~Y3    Y4~~Y4 
0.8444582 0.9968492 1.0012298 0.9868099 
```

Conversion to LMM coefficients:

```
list(sigma = sqrt(omega.lvm + lambda.lvm^2 * tau.lvm),
     rho = sqrt( lambda.lvm^2 * tau.lvm / (omega.lvm + lambda.lvm^2 * tau.
   lvm)))
```

```
$sigma
  Y1~~Y1   Y2~~Y2   Y3~~Y3   Y4~~Y4 
1.450466 1.392831 1.591194 1.544450 

$rho
             Y2~eta    Y3~eta    Y4~eta 
0.7737012 0.6972477 0.7775305 0.7657021 
```

Fit LMM:

```
rhoLVM <- function(p,time,...){
  R <- tcrossprod(p[time])
  diag(R) <- 1
  return(R)
}
myStruct <- CUSTOM(~time,
     FCT.sigma = function(p,time,X){p[time]},
     init.sigma = setNames(rep(1.45,n.time),paste0("sigma",1:n.time)),
     FCT.rho = rhoLVM,
     init.rho = setNames(rep(0.7,n.time),paste0("rho",1:n.time)))
e.lmmCUSTOM <- lmm(Y ~ time*age,
     repetition = ~time|id,
     structure  = myStruct, data = dfL.sim,
     method.fit = "ML")
logLik(e.lmmCUSTOM)
```

```
[1] -651.0478
```

We get exactly the same log-likelihood as the latent variable model. Export coefficient by type:

```
mu.lmm <- coef(e.lmmCUSTOM, effects = "mean")
sigma.lmm <- coef(e.lmmCUSTOM, effects = "variance")
rho.lmm <- coef(e.lmmCUSTOM, effects = "correlation")
list(mu = mu.lmm,
     sigma = sigma.lmm,
     rho = rho.lmm)
```

```
$mu
(Intercept)       time2       time3       time4         age   time2:age 
-0.18353676  0.03440620  0.16412896  0.32950080  0.95027744  0.10325886 
  time3:age   time4:age 
 0.01685224  0.08466026 

$sigma
  sigma1   sigma2   sigma3   sigma4 
1.450466 1.392831 1.591194 1.544450 

$rho
     rho1      rho2      rho3      rho4 
0.7737013 0.6972475 0.7775306 0.7657020 
```

Conversion to LVM coefficients:

```
list(lambda = rho.lmm*sigma.lmm/(rho.lmm[1]*sigma.lmm[1]),
     tau = rho.lmm[1]^2*sigma.lmm[1]^2,
     omega = sigma.lmm^2 * (1-rho.lmm^2))
```

```
$lambda
    rho1     rho2     rho3     rho4
1.000000 0.865375 1.102452 1.053786

$tau
    rho1
1.259395

$omega
   sigma1    sigma2    sigma3    sigma4
0.8444578 0.9968496 1.0012293 0.9868104
```