Chapter 1

Examples

The following chapter presents the Bayesian LMG implementation by two examples. Simulated data is used for the first example. Empirical data is used for the second example.

1.1 Simulated Data

We assume a simple model for the first example:

$$\begin{aligned}
&Y_i \sim \mathcal{N}(\beta_0 + x_1\beta_1 + x_2\beta_2 + x_3\beta_3 + x_4\beta_4,\, \sigma^2),\\
&\beta_1 = 0.5, , \beta_2 = 1,\, \beta_3 = 2,\, \beta_4 = 0,\, \sigma^2 = 1\\
&\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4 \sim \mathcal{N}(0, 1)
\end{aligned}$$

The values of the four predictors are sampled from a standard normal distribution. These values are then multiplied by the regression coefficients. A standard normal distributed error is added to obtain the dependent variable. Fifty observations were sampled. The data generating R-code **??** can be found in the Appendix.

The model is fitted using the `rstanarm` package (Stan Development Team, 2016) with the default priors for the regression and σ^2 parameters. The exact command can be found in R-code **??** These default priors are called 'weakly informative priors', because they take into account the order of magnitude of the variables by using the variance of the observed data. Information about these priors can be found in Stan Development Team (2017). A burn-in period of 20000, a sample size of 20000, and a thinning of 20 were chosen, resulting in a posterior sample size of 1000. The first few posterior samples are shown in Table 1.1.

For each posterior sample of the parameters, the R^2 value was calculated. The R^2 of the submodels was then calculated by the conditional variance formula for each posterior sample. The R-code is found in **??**. The resulting R^2 values of the first few posterior samples are shown in Table 1.2. The thinning is reasonable in this case to reduce the computational burden and to still obtain an appropriate posterior of the R^2 values (Link and Eaton, 2012).

The `hier.part` package was used to calculate the LMG value for each posterior sample. The LMG posteriors are shown in Table 1.3 . The independent component (I) represents the LMG value. The joint contribution (J) represents the difference from the independent component to the explained variance of the model containing only the predictor itself (T). Assuming stochastic or non-stochastic regressors has an influence on the uncertainty of the LMG values.

Table 1.1: Samples from the posterior distributions of the regression parameters

	x1	x2	x3	x4	sigma
sample 1	1.34	1.073	0.670	0.991	0.877
sample 2	1.07	1.396	0.937	0.999	1.055
sample 3	1.17	1.141	0.699	1.188	0.877
sample 4	1.29	1.064	0.640	1.136	1.012
sample 5	1.00	1.124	0.859	1.569	0.819
sample 6	0.99	1.102	0.599	1.127	1.035
sample 7	1.26	0.904	0.702	1.437	1.115
sample 8	1.22	1.128	0.552	1.421	0.924
sample 9	1.28	0.954	0.661	1.088	0.931
sample 10	1.09	1.057	0.865	1.475	0.984

Table 1.2: R^2 for all submodels for the first six posterior samples

	sample 1	sample 2	sample 3	sample 4	sample 5	sample 6
none	0.000	0.000	0.000	0.000	0.000	0.000
x1	0.351	0.213	0.253	0.296	0.155	0.196
x2	0.397	0.467	0.431	0.380	0.398	0.422
x3	0.081	0.117	0.078	0.067	0.100	0.057
x4	0.078	0.083	0.138	0.110	0.251	0.145
x1 x2	0.649	0.597	0.597	0.588	0.488	0.543
x1 x3	0.400	0.299	0.304	0.337	0.230	0.233
x1 x4	0.534	0.377	0.504	0.517	0.525	0.443
x2 x3	0.503	0.615	0.533	0.469	0.525	0.501
x2 x4	0.434	0.503	0.509	0.440	0.570	0.506
x3 x4	0.178	0.222	0.239	0.196	0.386	0.223
x1 x2 x3	0.721	0.716	0.672	0.648	0.593	0.601
x1 x2 x4	0.760	0.684	0.756	0.728	0.748	0.698
x1 x3 x4	0.601	0.484	0.575	0.574	0.630	0.497
x2 x3 x4	0.553	0.667	0.629	0.543	0.726	0.601
all	0.844	0.817	0.846	0.802	0.875	0.770

At first, non-stochastic regressors were assumed. The resulting LMG values and joint contributions with a 95% credible interval are shown in Table 1.3. An option to display the resulting LMG distribution is shown in Figure 1.1. Using the default weakly informative priors, the LMG distributions obtained from the Bayesian framework were very similar to the bootstrap confidence intervals, assuming non-stochastic predictors of the LMG estimates obtained from the `relaimpo` package, as shown in Table 1.4.

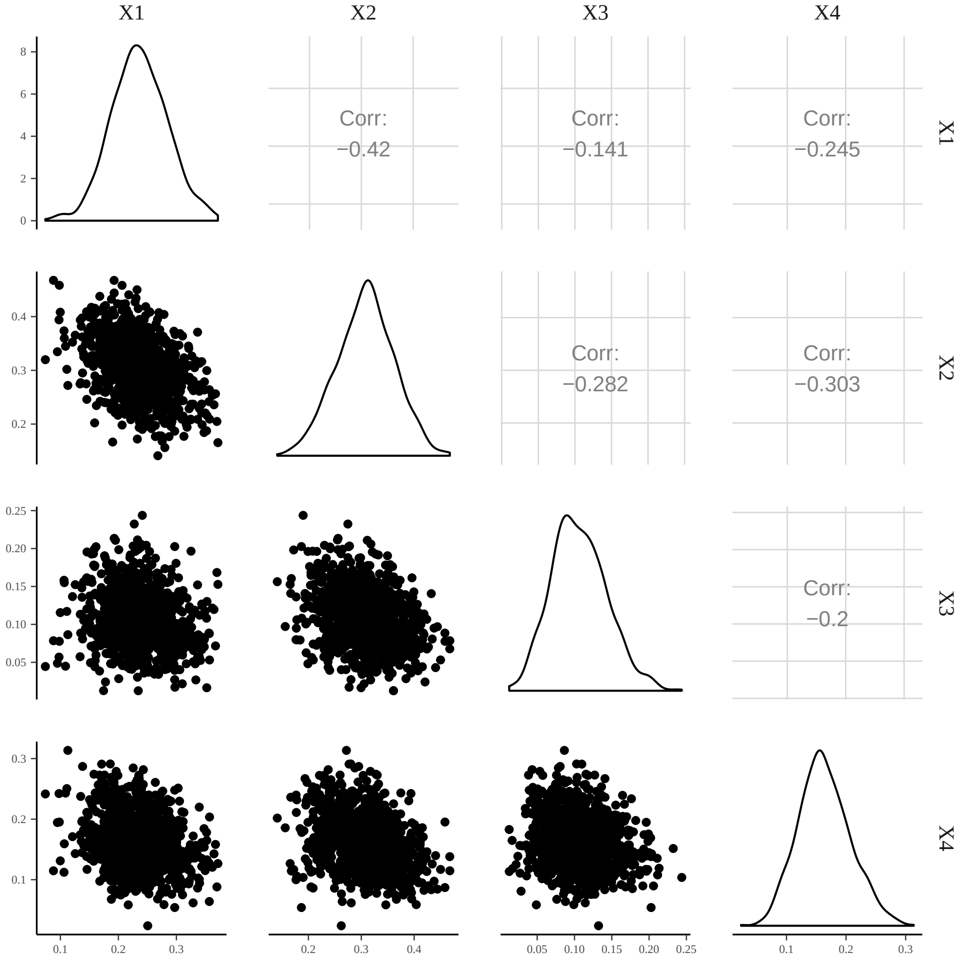

Figure 1.1: Posterior distribution of LMG values.

Table 1.3: Variance decomposition for non-stochastic predictors. I = LMG values, J = joint contribution, Total = total explained variance in one-predictor only model

Variable	**I**	**J**	**Total**
x1	0.235 (0.145, 0.337)	0.01 (-0.009, 0.029)	0.245 (0.15, 0.351)
x2	0.31 (0.199, 0.411)	0.066 (0.051, 0.079)	0.376 (0.256, 0.482)
x3	0.104 (0.041, 0.187)	-0.008 (-0.014, -0.002)	0.096 (0.035, 0.178)
x4	0.161 (0.085, 0.254)	-0.027 (-0.042, -0.013)	0.133 (0.064, 0.228)

Table 1.4: Variance decomposition for non-stochastic predictors. I = LMG values, J = joint contribution, Total = total explained variance in one-predictor only model

	LMG value (95%-CI)	
Variable	Relaimpo	**Bayesian framework**
x1	0.241 (0.152, 0.339)	0.235 (0.145, 0.337)
x2	0.318 (0.223, 0.431)	0.31 (0.199, 0.411)
x3	0.107 (0.046, 0.185)	0.104 (0.041, 0.187)
x4	0.166 (0.097, 0.262)	0.161 (0.085, 0.254)

In this example, we know, that the predictor values were sampled from a normal distribution. It would therefore be more reasonable to assume stochastic predictors. Under the assumption of weak exogeinity and conditional independence, the posterior distributions of the regression parameters $\boldsymbol{\beta}$ are valid for non-stochastic and stochastic predictors. However, the uncertainty about the LMG values needs to include the uncertainty about the covariance matrix. If we know the distribution of the predictors we can incorporate this information and obtain the posterior distribution of the covariance matrix. The package `Jags` was used for inference about the covariance matrix in a Bayesian way. As an alternative, non-parametric bootstrap was used for inference about the covariance matrix.

Using the bootstrap samples of the covariance matrix or samples from the posterior covariance matrix produced very similar LMG distributions. Bootstrap seems to be a valuable option for stochastic predictors when the distribution of the predictors is unknown. Even when the distribution is known, the difference seems to be tiny. A benefit of going the full Bayesian way is that we can also include prior knowledge about the covariance matrix. Using the default priors further produced very similar LMG distribution as using the non-parametric bootstrap option of the `relaimpo` package. Table 1.5 shows the LMG values of these approaches. For stochastic predictors, in contrast to non-stochastic predictors, the uncertainty about the covariance matrix is reflected in the larger credible intervals. Even when the exact regression parameters were known, there would a lot of uncertainty in the LMG values caused by the uncertainty about the covariance matrix.

Table 1.5: LMG values assuming stochastic predictors with 95% CI.

	Relaimpo	**Bayesian framework**	
Variable		nonparameteric bootstrap	covariance inference
x1	0.241 (0.125, 0.392)	0.242 (0.139, 0.356)	0.248 (0.137, 0.408)
x2	0.318 (0.187, 0.429)	0.288 (0.164, 0.423)	0.302 (0.186, 0.416)
x3	0.107 (0.022, 0.249)	0.092 (0.031, 0.186)	0.099 (0.033, 0.192)
x4	0.166 (0.075, 0.28)	0.176 (0.088, 0.301)	0.161 (0.058, 0.269)

Table 1.6: Variable description

Variable Name	Description
paragrap	scores on paragraph comprehension test
general	scores on general information test
sentence	scores on sentence completion test
wordc	scores on word classification test
wordm	scores on word meaning test

1.2 Empirical Data

In the following section, the Bayesian LMG implementation is applied on an empirical dataset containing test scores of pupils (N = 301) from a study by Holzinger and Swineford (1939) available in the R package `MBESS` (Kelley, 2017). This dataset was used in Nimon *et al.* (2008) to present commonality analysis, which is another variance decomposition technique. Scores from a paragraph comprehension test (paragrap) were predicted by four verbal tests: general-information (general), sentence-comprehension (sentence) , word-classification (wordc), and word-meaning (wordm) (Table 1.6).

The aim of the regression analysis was to determine the association between verbal ability and paragraph comprehension. An overview of the data is shown in Figure 1.3. The regression results are shown in Table 1.7). A novice researcher may wrongly conclude, that there is little association between the "non-significant" predictors (general information and word-classification) and paragraph comprehension. Given the other predictors are already included in the model, the predictors seem not to provide much information about the expected paragraph comprehension ability. However, it should not be concluded from this regression table, that the association between any of these "non-significant" predictors and the dependent variable is unimportant. As shown in Figure 1.3, the correlations between the predictors are rather high. The LMG metric may therefore provide new information about the importance of each predictor.

The Bayesian regression model was fitted in `rstanarm`. The default priors were used for the regression coefficients and the σ^2 parameter. A burn-in period of 20000, a sample size of 20000, and a thinning of 20 resulted in a posterior sample size of 1000. The first few posterior samples are shown in Table 1.8. The resulting R^2 of these posterior samples are shown in Table 1.9. The LMG values were calculated by using `hier.part`. The independent component (I), joint contribution (J), and total explained variance in a one-predictor model (T) are shown in Table 1.10. Sentence-comprehension and word-meaning seem to be the most important predictors by applying the LMG metric. However, none of the predictors seem to be unimportant. The joint contributions of each predictor were quite large.

For comparison purposes, the LMG metric was additionally calculated with the `relaimpo` package using parametric bootstrapping. The confidence intervals of `relaimpo` were almost identical to the credible intervals of the Bayesian framework (Table 1.11). Assuming stochastic or non-stochastic predictors resulted also in almost identical uncertainty estimates with such a large sample size (Table1.12).

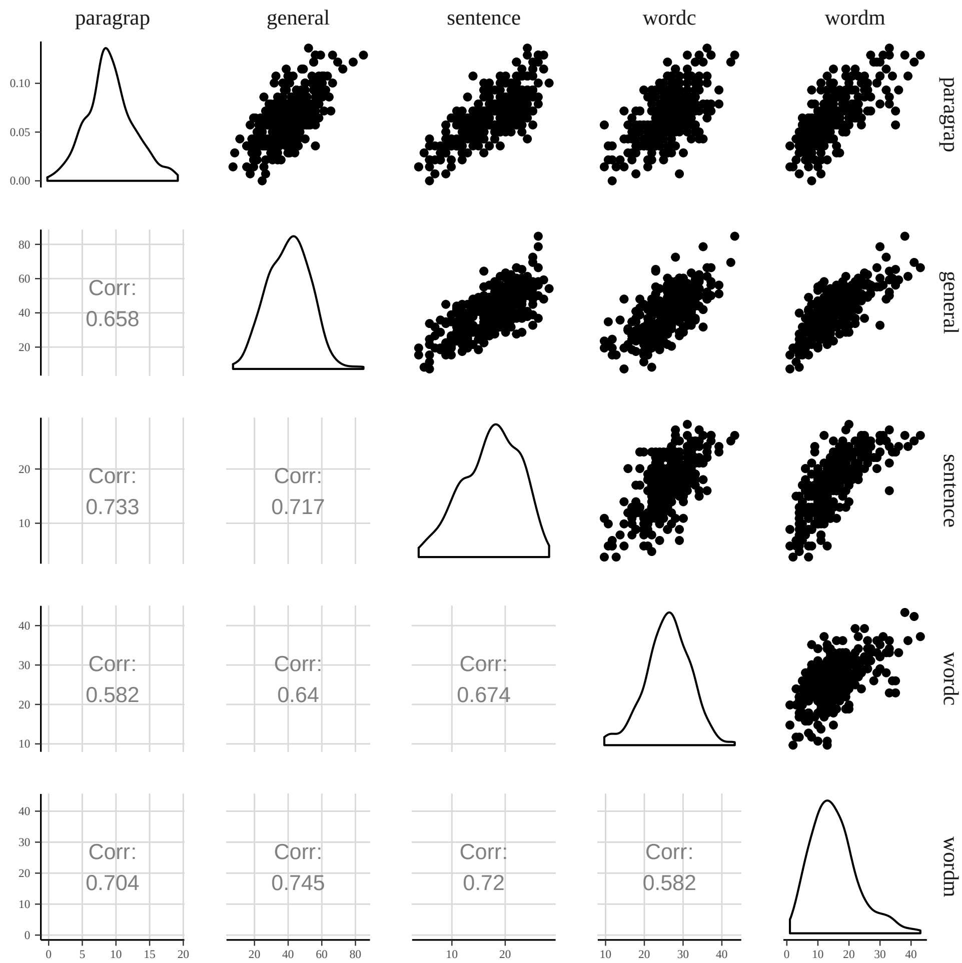

Figure 1.2: Test scores from Holzinger and Swineford's (1939) Study. N=301

Table 1.7: Regression of paragraph comprehension on verbal tests.

	Coefficient	95%-confidence interval	p-value
Intercept	0.071	from -1.17 to 1.31	0.91
general	0.03	from -0.00 to 0.06	0.084
sentence	0.26	from 0.18 to 0.34	$<$ 0.0001
wordc	0.047	from -0.01 to 0.11	0.14
wordm	0.14	from 0.08 to 0.19	$<$ 0.0001

Table 1.8: Samples from the posterior distributions of the regression parameters

	general	sentence	wordc	wordm	sigma
sample 1	0.035	0.251	0.047	0.133	2.06
sample 2	0.036	0.235	0.049	0.155	2.24
sample 3	0.050	0.268	0.036	0.150	2.15
sample 4	0.014	0.254	0.054	0.141	2.32
sample 5	0.026	0.318	0.022	0.130	2.25
sample 6	0.039	0.260	0.070	0.116	2.12
sample 7	0.063	0.239	0.039	0.091	2.27
sample 8	0.035	0.305	0.037	0.136	2.13
sample 9	0.051	0.284	0.009	0.137	2.19
sample 10	0.072	0.246	0.010	0.107	2.21

Table 1.9: R^2 for all submodels for the first six posterior samples

	sample 1	sample 2	sample 3	sample 4	sample 5	sample 6
none	0.000	0.000	0.000	0.000	0.000	0.000
general	0.456	0.442	0.494	0.373	0.416	0.460
sentence	0.551	0.518	0.571	0.488	0.552	0.555
wordc	0.350	0.335	0.355	0.310	0.314	0.373
wordm	0.512	0.510	0.544	0.452	0.479	0.493
general sentence	0.593	0.563	0.624	0.512	0.578	0.598
general wordc	0.499	0.481	0.531	0.419	0.453	0.513
general wordm	0.558	0.549	0.597	0.479	0.517	0.547
sentence wordc	0.567	0.534	0.584	0.501	0.559	0.576
sentence wordm	0.619	0.597	0.648	0.548	0.604	0.612
wordc wordm	0.559	0.550	0.586	0.494	0.517	0.555
general sentence wordc	0.597	0.567	0.627	0.517	0.579	0.605
general sentence wordm	0.628	0.606	0.661	0.550	0.607	0.623
general wordc wordm	0.579	0.568	0.613	0.503	0.534	0.577
sentence wordc wordm	0.625	0.603	0.653	0.553	0.605	0.623
all	0.631	0.609	0.663	0.554	0.608	0.630

Table 1.10: Variance decomposition for non-stochastic predictors. I = LMG values, J = joint contribution, Total = total explained variance in one-predictor only model

Variable	**I**	**J**	**Total**
general	0.13 (0.104, 0.159)	0.299 (0.258, 0.332)	0.429 (0.363, 0.489)
sentence	0.203 (0.166, 0.246)	0.327 (0.291, 0.356)	0.532 (0.467, 0.589)
wordc	0.095 (0.074, 0.127)	0.239 (0.201, 0.277)	0.334 (0.275, 0.401)
wordm	0.177 (0.141, 0.215)	0.315 (0.279, 0.344)	0.491 (0.423, 0.551)

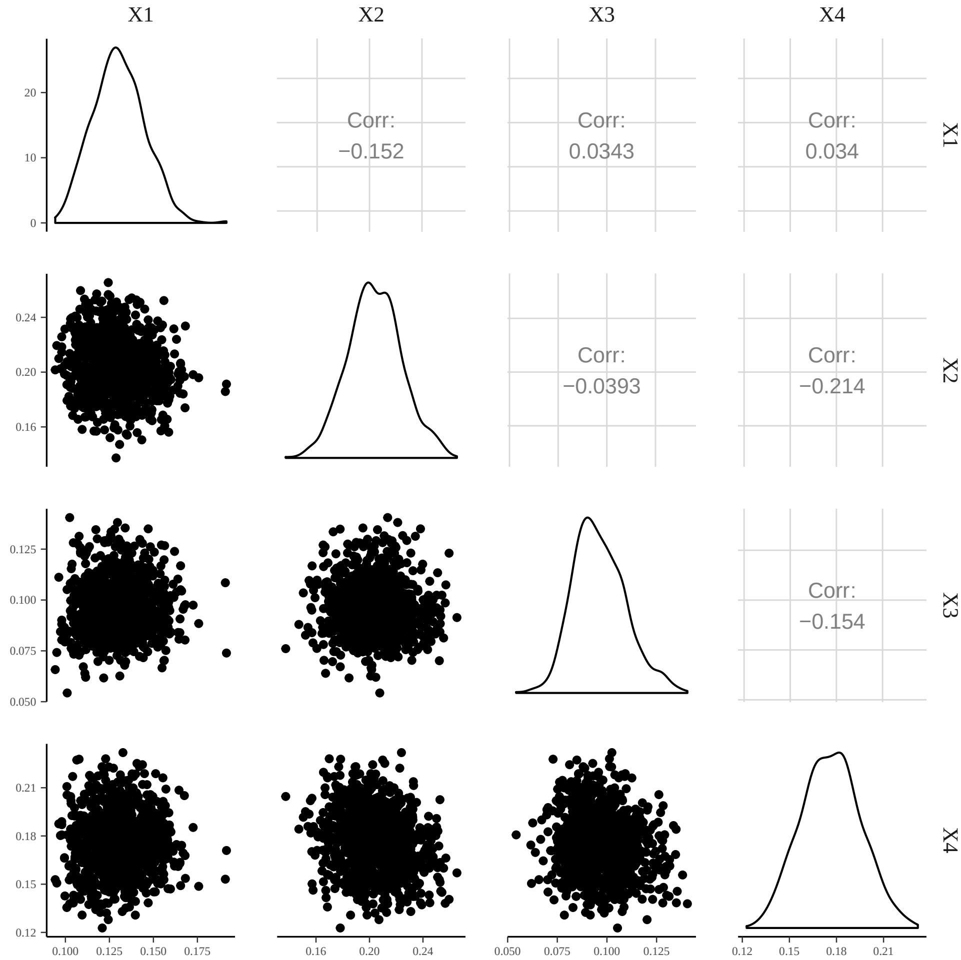

Figure 1.3: LMG posterior distribution of different verbal ability tests

Table 1.11: Variance decomposition for non-stochastic predictors. I = LMG values, J = joint contribution, Total = total explained variance in one-predictor only model

	LMG value (95%-CI)	
Variable	Relaimpo	**Bayesian framework**
general	0.131 (0.106, 0.162)	0.13 (0.104, 0.159)
sentence	0.206 (0.165, 0.248)	0.203 (0.166, 0.246)
wordc	0.096 (0.074, 0.127)	0.095 (0.074, 0.127)
wordm	0.178 (0.143, 0.217)	0.177 (0.141, 0.215)

Table 1.12: Variance decomposition for Stochastic predictors. I = LMG values, J = joint contribution, Total = total explained variance in one-predictor only model

	LMG value (95%-CI)	
Variable	Relaimpo	**Bayesian framework**
general	0.131 (0.106, 0.164)	0.127 (0.097, 0.16)
sentence	0.206 (0.17, 0.247)	0.202 (0.162, 0.246)
wordc	0.096 (0.075, 0.124)	0.094 (0.066, 0.128)
wordm	0.178 (0.144, 0.216)	0.175 (0.138, 0.216)

Bibliography

Holzinger, K. J. and Swineford, F. (1939). A study in factor analysis: The stability of a bi-factor solution. *Supplementary Educational Monographs*, **48**, 1–91. 5

Kelley, K. (2017). Mbess (version 4.0.0 and higher) [computer software and manual]. R package version 4.0.0 and higher. 5

Link, W. A. and Eaton, M. J. (2012). On thinning of chains in MCMC. *Methods in Ecology and Evolution*, **3**, 112–115. 1

Nimon, K., Lewis, M., Kane, R., and Haynes, R. M. (2008). An R package to compute commonality coefficients in the multiple regression case: An introduction to the package and a practical example. *Behavior Research Methods*, **40**, 457–466. 5

Stan Development Team (2016). rstanarm: Bayesian applied regression modeling via Stan. R package version 2.13.1. 1

Stan Development Team (2017). *Stan Modeling Language: User's Guide and Reference Manual*. 1