Take home exam

Gaussian hierarchical linear model

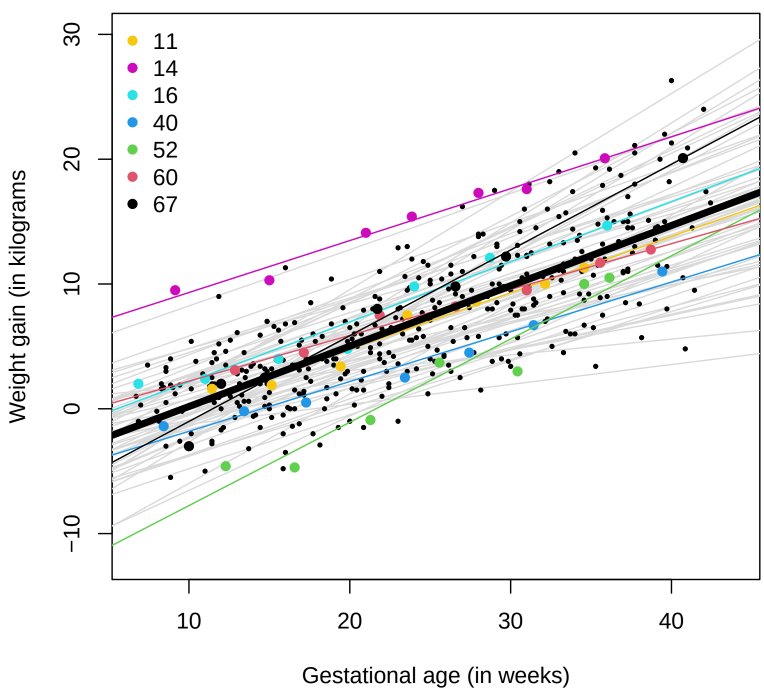
Advanced Bayesian Econometrics

PhD in Business Economics — Instructor: Hedibert Freitas Lopes

Start: 9am, February 18th, 2021. — End: 12pm, February 20th, 2021.

This is an individual exam. You are free to consult books, videos, short-courses or other studying material. You are not suppose though to consult your colleagues or any other individual to assist you throughout the examination.

Souza, in her 1999 PhD Thesis[1], considers a number of hierarchical and dynamic models to describe the nutritional pattern of pregnant women. The data we will analyze appear at the end of this document and in the file in our course webpage. Slides from a short-course entitled *Modelos Hierárquicos e Aplicações* is available here (you might wanna browse through pages 9-12 and 17-26). The data depicted in Figure 5.5 (page 161) of Gamerman and Lopes (2006) consist of the weight gain (in kilograms) of individual i, denoted here by y_{ij}, at her j^{th} visit at gestational age (in weeks), x_{ij}. The data set contains $I = 68$ pregnant women to visited $J = 7$ times the Instituto de Puericultura e Pediatria Martagão Gesteira (IPPMG/UFRJ). The plot below shows all pairs (x_{ij}, y_{ij}), all 68 OLS fits and the pooled fit. The pooled fit ignores the individual differences and simply runs an OLS regression based on $IJ = 68 \times 7$ pairs of points. In the plot we also highlight 7 randomly chosen patients. Notice that intercepts and slope vary widely.



[1] Aparecida D. P. Souza (1999) Approximate Methods in Bayesian Dynamic Hierarchical Models, unpublished Ph.D. Thesis, COPPE-UFRJ (in Portuguese).

HLM. One of the models she entertained is the Gaussian hierarchical linear model (G-HLM). The model is written, for $i = 1, \ldots, I$ and $j = 1, \ldots, J$, as a random intercept and random slope model (a.k.a as multilevel model or longitudinal model):

$$y_{ij}|\alpha_i, \beta_i, \sigma^2 \sim N(\alpha_i + \beta_i x_{ij}, \sigma_y^2) \quad (1)$$

$$(\alpha_i, \beta_i)'|\alpha, \beta \sim N((\alpha, \beta)', \text{diag}(\sigma_\alpha^2, \sigma_\beta^2)) \quad (2)$$

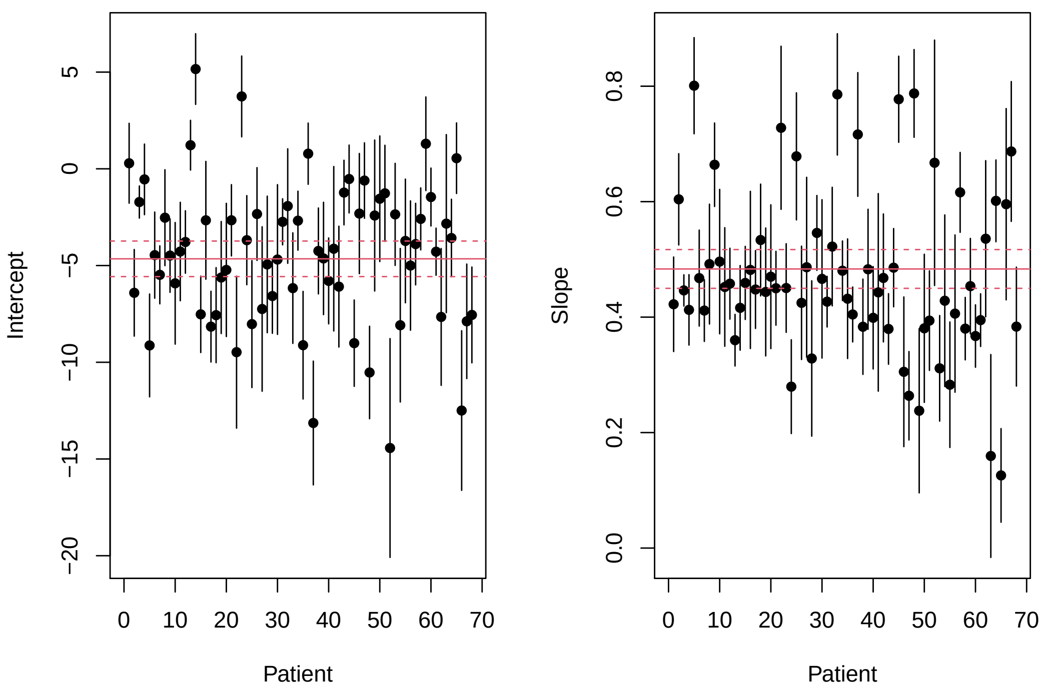
$$(\alpha, \beta)' \sim N((\alpha_0, \beta_0)', \text{diag}(\tau_\alpha^2, \tau_\beta^2)). \quad (3)$$

The variance parameters $\sigma_y^2, \sigma_\alpha^2$ and σ_β^2 are all *a priori* independent $IG(a_0, a_0)$, where $a_0 = 0.001$. Also, $\alpha_0 = \beta_0 = 0$ and $\tau_\alpha = \tau_\beta = 10$. In other words, at the very bottom of the hierarchical model the prior distributions are fairly non-informatives.

(a) Run one OLS regression (the pooled model), where $\alpha_i = \alpha$ and $\beta_i = \beta$ for $i = 1, \ldots, I$ in Equation (1). This will give you estimates of α, β and σ: $\widehat{\alpha}_{pool}$, $\widehat{\beta}_{pool}$ and $\widehat{\sigma}^2_{y,pool}$. Equations (2)-(3) are unimportant here.

(b) Run I OLS regressions, one per pregnant woman. Here you will collect I intercepts, I slopes, I standard deviations, i.e. $\widehat{\alpha}_i$, $\widehat{\beta}_i$ and $\widehat{\sigma}_{y,i}$, for $i = 1, \ldots, I$. Again, Equations (2)-(3) are irrelevant here.

The previous two exploratory exercises will give you a better understanding of what is gained or lost by moving from the pooled linear model to the I individual linear models with only J observations. For instance, you can compare the I intercepts $\widehat{\alpha}_i$s to $\widehat{\alpha}_{pool}$. Similarly with $\widehat{\beta}_i$s and $\widehat{\sigma}_{y,i}$s. The following figure plots intercepts and slopes for each patient's regressions along with their 95% confidence intervals. The red lines are the estimates based on the pooled data.



Posterior inference via Gibbs sampler

(c) Show that all full conditional distributions of the G-HLM are of known form. More precisely, they are all either bivariate Gaussian or Inverse-Gamma densities, rendering them suitable for a standard Gibbs Sampler algorithm.

Let me help you out by outlining a few facts about the full conditionals. But first, let us introduce some more notation. For $i = 1, \ldots, I$, $y_i = (y_{i1}, \ldots, y_{iJ})$, X_i a $J \times 2$ matrix whose j^{th} row is $(1, x_{ij})$, $\theta_i = (\alpha_i, \beta_i)$, $\theta = (\alpha, \beta)$ and $V = \text{diag}(\sigma^2_\alpha, \sigma^2_\beta)$. Even though α_i and β_i are *a priori* independent, that is not the case *a posteriori*.

1. $[(\alpha_i, \beta_i)]$
   The full conditional distribution of θ_i combines the likelihood of a Gaussian linear model from Equation (1) with the bivariate prior of θ_i from Equation (2):

$$
\begin{aligned}
p(\theta_i | y_i, X_i, \sigma^2_y, \alpha, \beta, \sigma^2_\alpha, \sigma^2_\beta) \quad &\propto \quad \exp\left\{-\frac{1}{2\sigma^2_y}(y_i - X_i\theta_i)'(y_i - X_i\theta_i)\right\} \\
&\times \quad \exp\left\{-0.5(\theta_i - \theta)'V^{-1}(\theta_i - \theta)\right\}. \qquad (4)
\end{aligned}
$$

   You can derive from here (piece of cake!) and show that it is a bivariate normal.

2. $[\sigma^2_y]$
   This should be the easiest one! The model variance σ^2_y affects ALL individuals and ALL hospital visits ($i = 1, \ldots, I$ and $j = 1, \ldots, J$). So the likelihood, as a function of σ^2_y is

$$
\mathcal{L}(\sigma^2_y | \cdots) \propto (\sigma^2_y)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2_y}\sum_{i=1}^{I}\sum_{j=1}^{J}(y_{ij} - \alpha_i - \beta_i x_{ij})^2\right\},
$$

   which is combined with $(\sigma^2_y)^{-(a_0+1)}\exp\{-a_0/\sigma^2_y\}$. Guess that their multiplication look like? Yes, an Inverse-Gamma distribution.

3. $[\alpha]$, $[\beta]$, $[\sigma^2_\alpha]$ and $[\sigma^2_\beta]$
   Notice that α only depends on $\alpha_1, \ldots, \alpha_I$ and that β only depends on $\beta_1, \ldots, \beta_I$. Neither depend on on the y_{ij}s and x_{ij}. The same applies for σ^2_α and σ^2_β. The full conditional distributions will be Gaussian for α and β and Inverse-Gamma for σ^2_α and σ^2_β.

(d) You are now ready to use your derivations from (c) in order to implement (I hope in R) the Gibbs sampler for the Gaussian hierarchical linear model and, ultimately, obtain posterior summaries. Compare the results from the Bayesian approach with those you found in (a) and (b).

patient	visit	gestational.age	weight.gain
1	1	8.571	3.3
1	2	13	6.1
1	3	17.571	8.5
1	4	20.857	7.9
1	5	24.571	11.8
1	6	35.429	14.8
1	7	40.143	17.2
2	1	14.429	2
2	2	18.714	4.3
2	3	23	8
2	4	27	11
2	5	31	12.2
2	6	37.429	15.6
2	7	37.857	16.5
3	1	13.429	4.5
3	2	18.857	6.8
3	3	23.143	8.1
3	4	34.143	13.5
3	5	37	15
3	6	38	15.2
3	7	39.286	15.8
4	1	12.571	5.5
4	2	17	5.5
4	3	21.429	8
4	4	32.429	13.2
4	5	34.286	13.9
4	6	38.571	15.1
4	7	40.571	16.2
5	1	13.714	2
5	2	23.571	9.5
5	3	27.714	12.2
5	4	32.429	18.2
5	5	35.286	19.3
5	6	39.571	22
5	7	41.857	24.4
6	1	11.286	1.8
6	2	14.857	2
6	3	21.286	4.5
6	4	25	7.1
6	5	30	9.5
6	6	34.429	12.6
6	7	39	13.5
7	1	13.143	0.2
7	2	17.143	1.4
7	3	23	3.6
7	4	28	5.8
7	5	31.429	8.3
7	6	36	9
7	7	38	10.2
8	1	10.429	3.8
8	2	14.429	3.4
8	3	17.857	5.4
8	4	21.857	8.8
8	5	25.714	10.4
8	6	30.429	12.3
8	7	36	15.3
9	1	8.857	1.9
9	2	14.429	5.9
9	3	17.714	6
9	4	23.429	10.6
9	5	28	14
9	6	37.714	21.1
9	7	38.429	21
10	1	12	0
10	2	15	1.3
10	3	19.857	3

```
10 4 24.286 6.4
10 5 26.714 8.3
10 6 31 10.8
10 7 37.286 11.2
11 1 11.429 1.6
11 2 15.143 1.9
11 3 19.429 3.4
11 4 23.571 7.5
11 5 27.857 8.6
11 6 32.143 10
11 7 34.571 11.3
12 1 9.143 1.2
12 2 14.857 3
12 3 19.143 4
12 4 26 8
12 5 28.571 9
12 6 34 12
12 7 39 14.5
13 1 10.143 5.4
13 2 15.571 6.3
13 3 26.286 10.8
13 4 29.286 11.2
13 5 33.286 13.3
13 6 37.286 15
13 7 37.571 14.7
14 1 9.143 9.5
14 2 15 10.3
14 3 21 14.1
14 4 23.857 15.4
14 5 28 17.3
14 6 31 17.6
14 7 35.857 20.1
15 1 12 -1.7
15 2 22.429 1.7
15 3 26.429 5.4
15 4 32.286 7.2
15 5 35.571 8.9
15 6 38 9.9
15 7 41.429 11.5
16 1 6.857 2
16 2 11 2.4
16 3 15.571 4
16 4 19.857 4.8
16 5 24 9.8
16 6 28.714 12.1
16 7 36 14.7
17 1 11.429 -2.8
17 2 15.857 -2
17 3 20.857 1.5
17 4 25.857 4.2
17 5 30.429 5.7
17 6 34.571 6.7
17 7 39.286 9.4
18 1 8.571 -3
18 2 14.143 -0.5
18 3 18.571 1.7
18 4 23.286 6
18 5 29.143 8.5
18 6 33.286 11
18 7 37.286 11
19 1 11.857 0.9
19 2 16.857 1.2
19 3 20.714 2.3
19 4 24.857 5
19 5 28.857 8
19 6 33.286 9.3
19 7 37 10.9
```

```
20 1 13.143 2
20 2 17.429 3.2
20 3 21.857 4.4
20 4 26.286 6.5
20 5 30.286 7.5
20 6 35.429 11.5
20 7 39.143 14.6
21 1 14 3.6
21 2 18.286 5.8
21 3 22.857 7.3
21 4 27 9
21 5 31.286 12.5
21 6 36.714 13.4
21 7 41.2 15.9
22 1 13.286 1.1
22 2 17.143 1.1
22 3 21.571 6.7
22 4 26.143 9.6
22 5 30.429 13.1
22 6 35.714 17.9
22 7 39.857 18.2
23 1 11.857 9
23 2 16 11.3
23 3 23 12.9
23 4 27 16.2
23 5 29 17.5
23 6 33 19
23 7 40 21.3
24 1 13.429 0.6
24 2 18.571 0.8
24 3 22.286 3
24 4 26.143 3.5
24 5 33.286 4.5
24 6 35.143 6.5
24 7 39.714 8
25 1 15.143 3.2
25 2 20.143 5.6
25 3 25.286 8.1
25 4 29.429 11.5
25 5 33.857 14.4
25 6 37.286 17
25 7 39.286 20
26 1 8.571 3
26 2 13.286 3.1
26 3 16.429 3.5
26 4 20.286 6
26 5 24.286 7.2
26 6 29.286 10
26 7 42.429 16.5
27 1 9.571 -0.2
27 2 14 -0.6
27 3 18.429 0
27 4 22.429 2
27 5 34.571 8.7
27 6 37 11
27 7 39 13.5
28 1 11.571 0.5
28 2 16.857 -1.2
28 3 20.143 1.6
28 4 24.143 3.2
28 5 28.857 3.8
28 6 33.714 6
28 7 38 8.4
29 1 14.714 0.9
29 2 19 3.5
29 3 23.286 7.2
29 4 29 9.5
```

```
29 5 33.286 11.6
29 6 39.286 14.4
29 7 39.571 15
30 1 14.714 3.2
30 2 19 4
30 3 23.714 5.5
30 4 28.571 8
30 5 31.143 9.8
30 6 35.143 10.7
30 7 37.571 14.5
31 1 10.143 1.6
31 2 17.286 4
31 3 20 6.5
31 4 24.5 7.7
31 5 30.714 10.5
31 6 35 12
31 7 39.857 14.3
32 1 11.857 5.2
32 2 19.714 7
32 3 23.714 9.7
32 4 28 13.8
32 5 33 15.4
32 6 35 16.4
32 7 40.286 19.1
33 1 10.857 2.8
33 2 15.286 6.6
33 3 21.857 11
33 4 24.857 11.5
33 5 31.143 18
33 6 34 20.5
33 7 40 26.3
34 1 11.429 3.7
34 2 15.429 3.9
34 3 20.429 6.8
34 4 31 12.3
34 5 36.429 15
34 6 38 15.6
34 7 40 16.5
35 1 11.429 -2.6
35 2 16 -3.5
35 3 19.286 -1.5
35 4 24.857 1.2
35 5 29.429 4
35 6 34 6
35 7 40.571 8.4
36 1 14.857 7
36 2 19.571 8.1
36 3 24.286 10.5
36 4 31.571 14.5
36 5 42.143 17.4
36 6 43 18.2
36 7 44 18.6
37 1 14.429 -1.5
37 2 20.286 0.3
37 3 25 4
37 4 30.143 8.4
37 5 33.143 10.3
37 6 37.286 14.5
37 7 40 15.5
38 1 8 -0.2
38 2 15 0
38 3 22.714 4.2
38 4 24.571 5.8
38 5 30.429 7.5
38 6 34.857 9.2
38 7 39 10.7
39 1 13 2.8
```

```
39 2 17.286 2.5
39 3 21.857 5.2
39 4 25.571 8.5
39 5 32.571 10.5
39 6 35.714 13.2
39 7 39 14.2
40 1 8.429 -1.4
40 2 13.429 -0.2
40 3 17.286 0.5
40 4 23.429 2.5
40 5 27.429 4.5
40 6 31.429 6.7
40 7 39.429 11
41 1 6.714 1
41 2 12.571 1
41 3 16.571 0
41 4 21.286 4.9
41 5 26.571 9.2
41 6 34.571 11.5
41 7 38.571 13
42 1 11.286 0.7
42 2 16.143 0.1
42 3 22.429 4.5
42 4 27.286 5.7
42 5 30.714 8
42 6 37.571 12.5
42 7 39.571 12.4
43 1 11.714 4
43 2 15.857 3.9
43 3 20.286 6
43 4 25.143 8.7
43 5 28.857 10
43 6 37.714 13
43 7 39.143 13.6
44 1 8.857 4
44 2 11.571 4.5
44 3 16 6.8
44 4 23.857 12
44 5 28.286 14
44 6 35.714 15.9
44 7 39.714 18.8
45 1 13.5 2.5
45 2 19.857 5.4
45 3 25 10.3
45 4 29.143 13.2
45 5 33.857 17.4
45 6 36.143 19.2
45 7 42 24
46 1 8.429 1.7
46 2 11.571 2
46 3 16.571 1.5
46 4 20.857 3
46 5 25.429 4.7
46 6 29.714 6
46 7 39.714 11.4
47 1 8.286 1.6
47 2 13.143 2
47 3 18.286 4
47 4 22.571 6.5
47 5 27.143 6.5
47 6 31.571 8.5
47 7 41.429 9.5
48 1 15 0.3
48 2 20.714 6
48 3 25 10
48 4 29.143 13
48 5 36.857 18.7
```

48 6 41 20.9
48 7 42 22.5
49 1 9.429 1.9
49 2 15.857 -0.5
49 3 20.429 1.5
49 4 25.286 3.8
49 5 29.857 3.8
49 6 35.714 7.5
49 7 41.286 7.4
50 1 7.429 3.5
50 2 11.429 2.5
50 3 16.571 3.5
50 4 20.471 5
50 5 27.429 8.1
50 6 35.857 12
50 7 39.571 15
51 1 12.286 4.6
51 2 16.857 5.1
51 3 22 6.4
51 4 27.571 9.5
51 5 32 10.6
51 6 37.286 14.5
51 7 41.429 15.1
52 1 12.286 -4.6
52 2 16.571 -4.7
52 3 21.286 -0.9
52 4 25.571 3.7
52 5 30.429 3
52 6 34.571 10
52 7 36.143 10.5
53 1 11.429 2
53 2 17.571 2.2
53 3 22.143 3.7
53 4 27.143 6.5
53 5 32.429 9
53 6 37.143 8.5
53 7 41.571 10.6
54 1 12.143 -1.5
54 2 16.429 -1.4
54 3 20.857 -1.5
54 4 25.143 2.8
54 5 29.286 5.7
54 6 33.429 6.2
54 7 39.143 8.7
55 1 14.714 0.2
55 2 19.143 1.2
55 3 23.571 3
55 4 27.714 4.5
55 5 32.143 7
55 6 38.143 5.7
55 7 40.714 7.8
56 1 8.143 -1
56 2 12.857 -0.7
56 3 16.286 0
56 4 21.857 4
56 5 25.143 7.5
56 6 30.857 8
56 7 40.714 10.5
57 1 8.286 2
57 2 15.714 5.4
57 3 26.286 11.5
57 4 30.571 14.2
57 5 37.714 20.5
57 6 39 20.1
57 7 40.143 20.8
58 1 13.571 3
58 2 18.571 3.8

```
58 3 23.143 6.5
58 4 31.429 8.8
58 5 34.429 11
58 6 37 11.5
58 7 39.143 12.3
59 1 14.857 7
59 2 18.857 10.4
59 3 23.571 13
59 4 30.571 15
59 5 32.286 16
59 6 37.714 18
59 7 38.286 18.7
60 1 12.857 3.1
60 2 17.143 4.5
60 3 21.857 7.5
60 4 26.571 8.2
60 5 31 9.5
60 6 35.571 11.7
60 7 38.714 12.8
61 1 6.857 -1
61 2 13.714 0.6
61 3 19.714 2.8
61 4 23.857 5.5
61 5 30 7.9
61 6 34.286 9.2
61 7 41.286 12
62 1 10.143 -2
62 2 15.143 -0.7
62 3 19.429 2.4
62 4 24.143 5.7
62 5 29.571 9.8
62 6 33 11.1
62 7 39.143 11.5
63 1 8.571 0.5
63 2 13 0.5
63 3 18.143 -2.9
63 4 23 -1
63 5 28.143 1.5
63 6 35.286 3.4
63 7 40.857 4.8
64 1 12.286 3.5
64 2 16.571 6.9
64 3 21.571 9
64 4 30.857 16
64 5 33.429 15.7
64 6 35 17.5
64 7 37.714 19.1
65 1 7 0.3
65 2 9 1.7
65 3 13.857 3.4
65 4 16.857 3.1
65 5 25.857 4.3
65 6 30 3.4
65 7 37 5.2
66 1 8.857 -5.5
66 2 11 -5
66 3 15.857 -4.8
66 4 20 -1
66 5 26.857 2.5
66 6 32.571 5
66 7 41.286 14.5
67 1 10 -3
67 2 12 2
67 3 14.714 2.5
67 4 21.714 8
67 5 26.571 9.8
67 6 29.714 12.2
```

```
67 7 40.714 20.1
68 1 9.429 -2.6
68 2 13.714 -3.2
68 3 17.714 -2
68 4 22 1
68 5 26.286 3
68 6 30.571 4.4
68 7 36.714 6.5
```