Here is the hierarchical model we will fit:

$$y_i = \beta_{0c[i]} + \beta_{1c[i]} X_{i1} + \beta_{2c[i]} X_{i2} + \beta_{3c[j]} X_{i3} + \eta_{c[i]} + \epsilon_i$$

The β_j are the coefficients, y_i is the response variable, X_{ij} are the predictor variables, $\eta_{c[i]}$ are the category error terms, one for each category containing the i data point. ϵ_i is the global error term.

$$\begin{aligned} \beta_j &\sim N(\mu_{jc}, \sigma_{jc}) \text{ for each } j \text{ and } c \\ \eta_c &\sim N(0, \sigma_\eta^2) \text{ for each } c \\ \epsilon_i &\sim N(0, \sigma_\epsilon^2) \end{aligned}$$

We have to fit the μ_{jc}'s, the σ_{jc}'s the σ_{η_c}'s and the σ_ϵ.