Hierarchical Bayesian Model

Our Study Group

UC Davis

2025-05-01

Big Picture: What Are We Modeling?

We are modeling **how antibody levels change over time** in response to infection, using data from multiple individuals and multiple **biomarkers** (10 antigen-isotype combinations, so (j = 1, 2, ..., 10)).

We want to:

- Understand the average pattern for each biomarker
- Allow each person's response to vary
- Share information across individuals to improve estimates

This is a perfect use case for a **hierarchical Bayesian model**.

Step 1: Individual-Level Parameters (Subject-Level)

Each person (i), for biomarker ($ j$), has their own unique set of parameters:

$$\theta_{ij} = \begin{bmatrix} y_{0,ij} \\ b_{0,ij} \\ \mu_{b,ij} \\ \mu_{y,ij} \\ \gamma_{ij} \\ \alpha_{ij} \\ \rho_{ij} \end{bmatrix}$$

These describe the antibody curve for person (i) and biomarker (j): the starting level, how fast it rises, peaks, and decays.

Step 2: Population-Level Parameters (Per Biomarker j)

Now we summarize how people typically behave for each **biomarker**:

$$\mu_j = \text{population mean vector for biomarker } j$$

This means:

- For biomarker (j), we believe the true average antibody trajectory is governed by parameters (μ_j).
- But we don't know (μ_j) — so we estimate it using data across all individuals.

Step 3: Hierarchical Modeling Structure

We assume each individual's parameter vector θ_{ij} is drawn from a multivariate normal distribution:

$$\theta_{ij} \sim \mathcal{N}(\mu_j, \Sigma_j)$$

- μ_j : the population-level mean for biomarker (j)
- Σ_j : $j \times j$ covariance matrix describing how the parameters co-vary

This is where the "**borrowing strength**" happens. Even if someone has sparse data, we can still make good inferences by **using the group-level pattern**.

Step 4: Priors on Population Means — "Hyperpriors"

But wait — since we are Bayesian, so we also need a prior belief about μ_j :

$$\mu_j \sim \mathcal{N}(\mu_{\text{hyp},j}, \Omega_{\text{hyp},j})$$

Where:

- $\mu_{\text{hyp},j}$: prior guess for the population mean (e.g., a vector of zeros))
- $\Omega_{\text{hyp},j}$: uncertainty about that guess (e.g., $100 \cdot I_7$ for weakly informative prior)

This is a **hyperprior**, because it's a prior on a prior-level parameter.

Step 5: Priors on Covariance — "Priors on Variability"

We also don't know how much individual parameters vary. So we assign a **Wishart prior** to the **inverse** covariance matrix:

$$\Sigma_j^{-1} \sim \mathcal{W}(\Omega_j, \nu_j)$$

- Ω_j : prior scale matrix (small variance across parameters, often $0.1 \cdot I_7$)
- ν_j : degrees of freedom

This tells the model how much we expect individuals to vary from the average for biomarker j.

Step 6: Measurement Error Model

Our observations are noisy! So we model the observed log-antibody levels $log(y_{obs,ij})$ like this:

$$\log(y_{\text{obs},ij}) \sim \mathcal{N}(\log(y_{\text{pred},ij}), \tau_j^{-1})$$

Where $\tau_j \sim \text{Gamma}(a_j, b_j)$ is a prior on measurement precision for biomarker j.

Step 7: Putting It All Together

The model is built hierarchically across five conceptual levels:

1. **Observed data**: log antibody concentrations from serum samples
2. **Individual-level parameters**: specific antibody dynamics for each subject-biomarker pair
3. **Population-level means**: average antibody parameters for each biomarker
4. **Hyperpriors on means**: our belief about the likely range of population means
5. **Priors on variability**: our belief about individual variation around those means

This structure lets us account for uncertainty at every level, while borrowing strength across subjects and biomarkers.

Summary of the Hierarchy

Let's stack it up top-down:

1. **Top Level**:
   - For each biomarker j, the true mean antibody trajectory parameters μ_j come from a prior $\mathcal{N}(\mu_{\text{hyp},j}, \Omega_{\text{hyp},j})$
2. **Middle Level**:
   - For each person i, their parameters $\theta_{ij} \sim \mathcal{N}(\mu_j, \Sigma_j)$
3. **Bottom Level**:
   - Their actual observed antibody levels are noisy measurements of predictions from θ_{ij}

RECAP: Where We Are

Step 3: Modeling Individuals

We say:

$$\theta_{ij} \sim \mathcal{N}(\mu_j, \Sigma_j)$$

This means:

- For biomarker j, each subject i has their own parameter vector θ_{ij}
- These vectors come from a **Normal distribution** centered at μ_j (the population mean for that biomarker)
- Σ_j is the covariance matrix capturing variation across individuals for that biomarker

But here's the catch: **we don't know** μ_j or Σ_j yet.

So How Do We Handle the Unknowns?

In **Bayesian modeling**, we treat unknowns as **random variables** too. So instead of fixing μ_j and Σ_j, we say:

> *"Let's estimate them, but we'll put a prior belief on them to guide the learning."*

This brings us to:

Step 4: Priors on μ_j

$$\mu_j \sim \mathcal{N}(\mu_{\text{hyp},j}, \Omega_{\text{hyp},j})$$

Explanation:

- μ_j: unknown population-level mean of the parameters for biomarker j
- We say:
  "We believe μ_j comes from another normal distribution"
  - *Centered at $\mu_{hyp,j}$ — a guess for what the mean might be*
  - *With spread $\Omega_{hyp,j}$ — how confident we are in that guess*

If we want to be very flexible, we make this prior **weakly informative**:

- Set $\mu_{\text{hyp},j} = 0$
- Set $\Omega_{\text{hyp},j} = 100 \cdot I_7$, where I_7 is the identity matrix (saying we are uncertain)

This is a **prior on a population-level parameter** — a "**hyperprior**".

Step 5: Priors on Σ_j

We also don't know how much individual parameters vary, so we say:

$$\Sigma_j^{-1} \sim \mathcal{W}(\Omega_j, \nu_j)$$

This is a **Wishart prior** on the **precision matrix** (inverse of covariance). Why?

- In multivariate stats, it's common to use the Wishart distribution as a prior for covariance matrices
- Ω_j: the scale (like the average covariance we expect)
- ν_j: degrees of freedom (how confident we are)

If we want to be uninformative, we might say:

- $\Omega_j = 0.1 \cdot I_7$
- $\nu_j = 8$

That allows a wide range of possible covariance matrices.

Summary of Why Priors Show Up

Priors appear at step 4 and 5 because we are now **modeling the parameters themselves**.

In Bayesian statistics:

- Every unknown quantity is treated as a random variable
- Every random variable must have a probability distribution
- That's what **priors** are

They let us encode our beliefs, and importantly, they let us **regularize the model** so it doesn't overfit sparse or noisy data.