Integrating over the random effects in a Bayesian hierarchical model

(first draft) by Jesse

I. Some lemmas

Completing the square in matrix form

If $\boldsymbol{A}$ is symmetric and invertible,

$$\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} - 2\boldsymbol{b}'\boldsymbol{y} = (\boldsymbol{y} - \boldsymbol{A}^{-1}\boldsymbol{b})'A(\boldsymbol{y} - \boldsymbol{A}^{-1}\boldsymbol{b}) - \boldsymbol{b}'\boldsymbol{A}^{-1}\boldsymbol{b},$$

or equivalently,

$$\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} - 2\boldsymbol{b}'\boldsymbol{y} + \boldsymbol{b}'\boldsymbol{A}^{-1}\boldsymbol{b} = (\boldsymbol{y} - \boldsymbol{A}^{-1}\boldsymbol{b})'A(\boldsymbol{y} - \boldsymbol{A}^{-1}\boldsymbol{b}).$$

Proof

Starting with the right hand side,

$$(\boldsymbol{y} - \boldsymbol{A}^{-1}\boldsymbol{b})'A(\boldsymbol{y} - \boldsymbol{A}^{-1}\boldsymbol{b}) - \boldsymbol{b}'\boldsymbol{A}^{-1}\boldsymbol{b}$$
$$= \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{A}\boldsymbol{A}^{-1}\boldsymbol{b} - \boldsymbol{b}'(\boldsymbol{A}^{-1})'\boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}'(\boldsymbol{A}^{-1})'\boldsymbol{A}\boldsymbol{A}^{-1}\boldsymbol{b} - \boldsymbol{b}'\boldsymbol{A}^{-1}\boldsymbol{b}$$

and simplifying and using the symmetry of $\boldsymbol{A}$, we have

$$= \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} - 2\boldsymbol{b}'\boldsymbol{y}.\ \blacksquare$$

Proposition

If random vector $\boldsymbol{y}$ has density $f(\boldsymbol{y}) \propto \exp\left\{-\frac{1}{2}Q(\boldsymbol{y})\right\}$,
where $Q(\boldsymbol{y}) = (\boldsymbol{y} - \boldsymbol{m})'\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{m})$,
then $\boldsymbol{y} \sim N_k(\boldsymbol{m}, \boldsymbol{V})$.

Proof

Suppose $f(\boldsymbol{y}) = C\exp\left\{-\frac{1}{2}Q(\boldsymbol{y})\right\}$ for some constant $C \neq 0$.

Then $f(\boldsymbol{y}) = C\left[(2\pi)^{-\frac{k}{2}}|\boldsymbol{V}|^{-\frac{1}{2}}\right]^{-1}(2\pi)^{-\frac{k}{2}}|V|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}Q(\boldsymbol{y})\right\}$.

The quantity $(2\pi)^{-\frac{k}{2}}|\boldsymbol{V}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}Q(\boldsymbol{y})\right\}$ is a multivariate normal density,

so integrating both sides wrt $\boldsymbol{y}$ gives $1 = C\left[(2\pi)^{-\frac{k}{2}}|\boldsymbol{V}|^{-\frac{1}{2}}\right]^{-1}$,

so $C = (2\pi)^{-\frac{k}{2}}|\boldsymbol{V}|^{-\frac{1}{2}}$, so $f(\boldsymbol{y}) = (2\pi)^{-\frac{k}{2}}|V|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}Q(\boldsymbol{y})\right\}$,

which is the $N_k(\boldsymbol{m}, \boldsymbol{V})$ density. $\blacksquare$

Proposition

If $f(\boldsymbol{y}) \propto \exp\left\{-\frac{1}{2}Q(\boldsymbol{y})\right\}$,
where $Q(\boldsymbol{y}) = \boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{y} - 2\boldsymbol{m}'\boldsymbol{V}^{-1}\boldsymbol{y}$,
then $\boldsymbol{y} \sim N_k(\boldsymbol{m}, \boldsymbol{V})$.

Proof

Rewrite $Q(\boldsymbol{y})$ as $(\boldsymbol{y}-\boldsymbol{m})'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{m}) - \boldsymbol{m}'\boldsymbol{V}^{-1}\boldsymbol{m}$, so that we have

$$f(\boldsymbol{y}) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{m})'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{m})\right\}\exp\left\{-\frac{1}{2}\boldsymbol{m}'\boldsymbol{V}^{-1}\boldsymbol{m}\right\}$$

$$\propto \exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{m})'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{m})\right\},$$

and the conclusion follows by the previous result. ■

Proposition

If $f(\boldsymbol{y}) \propto \exp\left\{-\frac{1}{2}Q(\boldsymbol{y})\right\}$,
where $Q(\boldsymbol{y}) = \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} - 2\boldsymbol{b}'\boldsymbol{y}$, and $\boldsymbol{A}$ is pd,
then $\boldsymbol{y} \sim N_k(\boldsymbol{A}^{-1}\boldsymbol{b}, \boldsymbol{A}^{-1})$.

Proof

Suppose $Q(\boldsymbol{y}) = \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} - 2\boldsymbol{b}'\boldsymbol{y}$,
and also $Q(\boldsymbol{y}) = \boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{y} - 2\boldsymbol{m}'\boldsymbol{V}^{-1}\boldsymbol{y}$ for some $\boldsymbol{m}, \boldsymbol{V}$.
Then $\boldsymbol{A} = \boldsymbol{V}^{-1}$, so $\boldsymbol{V} = \boldsymbol{A}^{-1}$.
Also $\boldsymbol{b}' = \boldsymbol{m}'\boldsymbol{V}^{-1}$, so $\boldsymbol{b} = \boldsymbol{V}^{-1}\boldsymbol{m}$, so $\boldsymbol{m} = \boldsymbol{V}\boldsymbol{b} = \boldsymbol{A}^{-1}\boldsymbol{b}$.
Therefore by the previous proposition, $\boldsymbol{y} \sim N_k(\boldsymbol{A}^{-1}\boldsymbol{b}, \boldsymbol{A}^{-1})$. ■

II. Model specification in matrix form

For y_{ijk} and ϵ_{ijk}, i is subject, j is timepoint, k is outcome.
In this model we have J timepoints and 3 outcomes.
For x_{ip}, α_{ip}, and β_{ip}, i is subject, p is index.
Ignore the effect of diagnostic group for now.

The matrix equation for subject i is $\boldsymbol{y_i} = \boldsymbol{X_i}\boldsymbol{\alpha_i} + \boldsymbol{W_i}\boldsymbol{\beta_i} + \boldsymbol{\epsilon_i}$, or

$$\begin{bmatrix} y_{i11} \\ y_{i12} \\ y_{i13} \\ \cdots \\ y_{iJ1} \\ y_{iJ2} \\ y_{iJ3} \end{bmatrix} = \begin{bmatrix} x_{i1} & x_{i2} & & & & \\ & & x_{i1} & x_{i2} & & \\ & & & & x_{i1} & x_{i2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{i1} & x_{i2} & & & & \\ & & x_{i1} & x_{i2} & & \\ & & & & x_{i1} & x_{i2} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \alpha_6 \end{bmatrix} + \begin{bmatrix} 1 & t_{i1} & & & & \\ & & 1 & t_{i1} & & \\ & & & & 1 & t_{i1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & t_{iJ} & & & & \\ & & 1 & t_{iJ} & & \\ & & & & 1 & t_{iJ} \end{bmatrix} \begin{bmatrix} \beta_{i1} \\ \beta_{i2} \\ \beta_{i3} \\ \beta_{i4} \\ \beta_{i5} \\ \beta_{i6} \end{bmatrix} + \begin{bmatrix} \epsilon_{i11} \\ \epsilon_{i12} \\ \epsilon_{i13} \\ \cdots \\ \epsilon_{iJ1} \\ \epsilon_{iJ2} \\ \epsilon_{iJ3} \end{bmatrix},$$

with $\boldsymbol{\epsilon_i} \sim N_{3J}(\boldsymbol{0}, \boldsymbol{V})$, and with $\boldsymbol{V} = \boldsymbol{I_J} \otimes \boldsymbol{\Sigma}$,
where $\boldsymbol{\Sigma}$ is the (3×3) covariance matrix of the outcomes,
and with $\boldsymbol{\beta_i} \sim N_6(\boldsymbol{\mu}, \boldsymbol{\Omega})$.

For compactness, we might write

$$\boldsymbol{X_i} = \boldsymbol{1_J} \otimes \boldsymbol{I_3} \otimes \boldsymbol{x_i'}$$

$$\boldsymbol{W_i} = \begin{bmatrix} \boldsymbol{I_3} \otimes [1 \quad t_{i1}] \\ \cdots \\ \boldsymbol{I_3} \otimes [1 \quad t_{iJ}] \end{bmatrix}.$$

III. Integrating out $\boldsymbol{\beta}$

Assume $\boldsymbol{X}, \boldsymbol{\alpha}, \boldsymbol{W}, \boldsymbol{\Sigma}, \boldsymbol{\mu}, \boldsymbol{\Omega}$ are known. To integrate out $\boldsymbol{\beta}$, use the relationship

$$f(\boldsymbol{y}) = \int f(\boldsymbol{y}|\boldsymbol{\beta}) f(\boldsymbol{\beta}) d\boldsymbol{\beta}.$$

The densities are

$$f(\boldsymbol{y}|\boldsymbol{\beta}) \propto \exp\left\{-\tfrac{1}{2} Q_1(\boldsymbol{y}, \boldsymbol{\beta})\right\} \text{ and } f(\boldsymbol{\beta}) \propto \exp\left\{-\tfrac{1}{2} Q_2(\boldsymbol{\beta})\right\},$$

where (dropping subscripts)

$$Q_1(\boldsymbol{y}, \boldsymbol{\beta}) = (\boldsymbol{y} - \boldsymbol{X\alpha} - \boldsymbol{W\beta})'\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X\alpha} - \boldsymbol{W\beta})$$
$$Q_2(\boldsymbol{\beta}) = (\boldsymbol{\beta} - \boldsymbol{\mu})'\boldsymbol{\Omega}^{-1}(\boldsymbol{\beta} - \boldsymbol{\mu})$$

so the joint density is

$$\propto \exp\left\{-\tfrac{1}{2} Q(\boldsymbol{y}, \boldsymbol{\beta})\right\}$$

where $Q(\boldsymbol{y}, \boldsymbol{\beta}) = Q_1(\boldsymbol{y}, \boldsymbol{\beta}) + Q_2(\boldsymbol{\beta})$.

We add the quadratic forms and refactor as follows:

$$\begin{aligned}
&Q_1(\boldsymbol{y}, \boldsymbol{\beta}) + Q_2(\boldsymbol{\beta}) \\
&= (\boldsymbol{y} - \boldsymbol{X\alpha} - \boldsymbol{W\beta})'\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X\alpha} - \boldsymbol{W\beta}) + (\boldsymbol{\beta} - \boldsymbol{\mu})'\boldsymbol{\Omega}^{-1}(\boldsymbol{\beta} - \boldsymbol{\mu}) \\
&= \boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{X\alpha} - \boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{W\beta} - \boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} + \boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\alpha} + \boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{W\beta} \\
&\quad - \boldsymbol{\beta}'\boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{y} + \boldsymbol{\beta}'\boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{X\alpha} + \boldsymbol{\beta}'\boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{W\beta} + \boldsymbol{\beta}'\boldsymbol{\Omega}^{-1}\boldsymbol{\beta} - 2\boldsymbol{\mu}'\boldsymbol{\Omega}^{-1}\boldsymbol{\beta} + \boldsymbol{\mu}'\boldsymbol{\Omega}^{-1}\boldsymbol{\mu} \\
&= \boldsymbol{y}'[\boldsymbol{V}^{-1}]\boldsymbol{y} - 2[\boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{V}^{-1}]\boldsymbol{y} \\
&\quad -2\boldsymbol{\beta}'[\boldsymbol{W}'\boldsymbol{V}^{-1}]\boldsymbol{y} \\
&\quad + \boldsymbol{\beta}'[\boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{W} + \boldsymbol{\Omega}^{-1}]\boldsymbol{\beta} - 2[\boldsymbol{\mu}'\boldsymbol{\Omega}^{-1} - \boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{W}]\boldsymbol{\beta} \\
&\quad + \boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\alpha} + \boldsymbol{\mu}'\boldsymbol{\Omega}^{-1}\boldsymbol{\mu} \\
&= \boldsymbol{\beta}'[\boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{W} + \boldsymbol{\Omega}^{-1}]\boldsymbol{\beta} - 2[\boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{W} + \boldsymbol{\mu}'\boldsymbol{\Omega}^{-1} - \boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{W}]\boldsymbol{\beta} \\
&\quad + \boldsymbol{y}'[\boldsymbol{V}^{-1}]\boldsymbol{y} - 2[\boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{V}^{-1}]\boldsymbol{y} + \boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\alpha} + \boldsymbol{\mu}'\boldsymbol{\Omega}^{-1}\boldsymbol{\mu}.
\end{aligned}$$

Let $\boldsymbol{A} = \boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{W} + \boldsymbol{\Omega}^{-1}$, and let $\boldsymbol{b}' = \boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{W} + \boldsymbol{\mu}'\boldsymbol{\Omega}^{-1} - \boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{W}$, noting that $\boldsymbol{b}$ depends on $\boldsymbol{y}$, and we have

$$\begin{aligned}
&Q_1(\boldsymbol{y}, \boldsymbol{\beta}) + Q_2(\boldsymbol{\beta}) \\
&= \boldsymbol{\beta}'\boldsymbol{A\beta} - \boldsymbol{2b}'\boldsymbol{\beta} + \boldsymbol{b}'\boldsymbol{A}^{-1}\boldsymbol{b} \\
&\quad + \boldsymbol{y}'[\boldsymbol{V}^{-1}]\boldsymbol{y} - 2[\boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{V}^{-1}]\boldsymbol{y} + \boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\alpha} + \boldsymbol{\mu}'\boldsymbol{\Omega}^{-1}\boldsymbol{\mu} - \boldsymbol{b}'\boldsymbol{A}^{-1}\boldsymbol{b} \\
&= (\boldsymbol{\beta} - \boldsymbol{A}'\boldsymbol{b})'\boldsymbol{A}(\boldsymbol{\beta} - \boldsymbol{A}^{-1}\boldsymbol{b}) \\
&\quad + \boldsymbol{y}'[\boldsymbol{V}^{-1}]\boldsymbol{y} - 2[\boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{V}^{-1}]\boldsymbol{y} + \boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\alpha} + \boldsymbol{\mu}'\boldsymbol{\Omega}^{-1}\boldsymbol{\mu} - \boldsymbol{b}'\boldsymbol{A}^{-1}\boldsymbol{b}.
\end{aligned}$$

Now if we integrate over $\boldsymbol{\beta}$, the kernel

$$\propto \exp\left\{-\tfrac{1}{2}\left(\boldsymbol{\beta} - \boldsymbol{A}^{-1'}\boldsymbol{b}\right)'\boldsymbol{A}(\boldsymbol{\beta} - \boldsymbol{A}^{-1}\boldsymbol{b})\right\}$$

becomes 1, and we have

$$Q(\boldsymbol{y}) = \boldsymbol{y}'[\boldsymbol{V}^{-1}]\boldsymbol{y} - 2[\boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{V}^{-1}]\boldsymbol{y} + \boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\alpha} + \boldsymbol{\mu}'\boldsymbol{\Omega}^{-1}\boldsymbol{\mu} - \boldsymbol{b}'\boldsymbol{A}^{-1}\boldsymbol{b}.$$

At this point, WLOG we let $\boldsymbol{X\alpha} = \boldsymbol{0}$, remembering to add it to the mean of $\boldsymbol{y}$ at the end:

$$\boldsymbol{b}' = \boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{W} + \boldsymbol{\mu}'\boldsymbol{\Omega}^{-1}$$
$$Q(\boldsymbol{y}) = \boldsymbol{y}'[\boldsymbol{V}^{-1}]\boldsymbol{y} + \boldsymbol{\mu}'\boldsymbol{\Omega}^{-1}\boldsymbol{\mu} - (\boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{W} + \boldsymbol{\mu}'\boldsymbol{\Omega}^{-1})\boldsymbol{A}^{-1}(\boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{y} + \boldsymbol{\Omega}^{-1}\boldsymbol{\mu})$$

$$\begin{aligned}
&= \boldsymbol{y}'[\boldsymbol{V}^{-1}]\boldsymbol{y} + \boldsymbol{\mu}'\boldsymbol{\Omega}^{-1}\boldsymbol{\mu} - \boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{A}^{-1}\boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{y} - 2\boldsymbol{\mu}'\boldsymbol{\Omega}^{-1}\boldsymbol{A}^{-1}\boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{y} \\
&\quad + \boldsymbol{\mu}'\boldsymbol{\Omega}^{-1}\boldsymbol{A}^{-1}\boldsymbol{\Omega}^{-1}\boldsymbol{\mu} \\
&= \boldsymbol{y}'[\boldsymbol{V}^{-1} + \boldsymbol{V}^{-1}\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{W} + \boldsymbol{\Omega}^{-1})^{-1}\boldsymbol{W}'\boldsymbol{V}^{-1}]\boldsymbol{y} \\
&\quad - 2[\boldsymbol{\mu}'\boldsymbol{\Omega}^{-1}(\boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{W} + \boldsymbol{\Omega}^{-1})^{-1}\boldsymbol{W}'\boldsymbol{V}^{-1}]\boldsymbol{y} \\
&\quad + \boldsymbol{\mu}'[\boldsymbol{\Omega}^{-1} + \boldsymbol{\Omega}^{-1}(\boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{W} + \boldsymbol{\Omega}^{-1})^{-1}\boldsymbol{\Omega}^{-1}]\boldsymbol{\mu}.
\end{aligned}$$

Let $\boldsymbol{C} = \boldsymbol{V}^{-1} + \boldsymbol{V}^{-1}\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{W} + \boldsymbol{\Omega}^{-1})^{-1}\boldsymbol{W}'\boldsymbol{V}^{-1}$,
and let $\boldsymbol{d}' = \boldsymbol{\mu}'\boldsymbol{\Omega}^{-1}(\boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{W} + \boldsymbol{\Omega}^{-1})^{-1}\boldsymbol{W}'\boldsymbol{V}^{-1}$.
Then we have $Q(\boldsymbol{y}) = (\boldsymbol{y} - \boldsymbol{C}^{-1}\boldsymbol{d})'\boldsymbol{C}(\boldsymbol{y} - \boldsymbol{C}^{-1}\boldsymbol{d}) + [\text{other terms}]$,
and so $\boldsymbol{y} \sim \boldsymbol{N}(\boldsymbol{C}^{-1}\boldsymbol{d}, \boldsymbol{C}^{-1})$.
Adding back $\boldsymbol{X\alpha}$, we finally have

$$\begin{aligned}
&\boldsymbol{y} \sim N_{3J}(\boldsymbol{X\alpha} + \boldsymbol{C}^{-1}\boldsymbol{d}, \boldsymbol{C}^{-1}) \\
&\boldsymbol{C} = \boldsymbol{V}^{-1} + \boldsymbol{V}^{-1}\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{W} + \boldsymbol{\Omega}^{-1})^{-1}\boldsymbol{W}'\boldsymbol{V}^{-1} \\
&\boldsymbol{d} = \boldsymbol{V}^{-1}\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{W} + \boldsymbol{\Omega}^{-1})^{-1}\boldsymbol{\Omega}^{-1}\boldsymbol{\mu}. \blacksquare
\end{aligned}$$

Quick sanity check: do we see a $\boldsymbol{\beta}$ anywhere? No.
Another one: do dimensions match?
$\boldsymbol{X}$ is $(3J \times 6)$, $\boldsymbol{\alpha}$ is (6×1), so $\boldsymbol{X\alpha}$ is $(3J \times 1)$
$\boldsymbol{V}^{-1}$ is $(3J \times 3J)$, $\boldsymbol{W}$ is $(3J \times 6)$, so $\boldsymbol{C}$ and $\boldsymbol{C}^{-1}$ are $(3J \times 3J)$
$\boldsymbol{d}$ is $(3J \times 1)$, so $\boldsymbol{C}^{-1}\boldsymbol{d}$ is $(3J \times 1)$
so they do match.

IV. Another method

This time we make the means of the random intercept and slope fixed, part of α, so that each $\boldsymbol{\beta_i}$ has mean 0.
The matrix equation for subject i is $\boldsymbol{y_i} = \boldsymbol{X_i}\boldsymbol{\alpha_i} + \boldsymbol{W_i}\boldsymbol{\beta_i} + \boldsymbol{\epsilon_i}$, or

$$\begin{bmatrix} y_{i11} \\ y_{i12} \\ y_{i13} \\ \cdots \\ y_{iJ1} \\ y_{iJ2} \\ y_{iJ3} \end{bmatrix} = \begin{bmatrix} x_{i1} & x_{i2} & x_{i3} & x_{i4} & & & & & & & & \\ & & & & x_{i1} & x_{i2} & x_{i3} & x_{i4} & & & & \\ & & & & & & & & x_{i1} & x_{i2} & x_{i3} & x_{i4} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{i1} & x_{i2} & x_{i3} & x_{i4} & & & & & & & & \\ & & & & x_{i1} & x_{i2} & x_{i3} & x_{i4} & & & & \\ & & & & & & & & x_{i1} & x_{i2} & x_{i3} & x_{i4} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \alpha_6 \\ \alpha_7 \\ \alpha_8 \\ \alpha_9 \\ \alpha_{10} \\ \alpha_{11} \\ \alpha_{12} \end{bmatrix}$$

$$+ \begin{bmatrix} 1 & t_{i1} & & & & \\ & & 1 & t_{i1} & & \\ & & & & 1 & t_{i1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & t_{iJ} & & & & \\ & & 1 & t_{iJ} & & \\ & & & & 1 & t_{iJ} \end{bmatrix} \begin{bmatrix} \beta_{i1} \\ \beta_{i2} \\ \beta_{i3} \\ \beta_{i4} \\ \beta_{i5} \\ \beta_{i6} \end{bmatrix} + \begin{bmatrix} \epsilon_{i11} \\ \epsilon_{i12} \\ \epsilon_{i13} \\ \cdots \\ \epsilon_{iJ1} \\ \epsilon_{iJ2} \\ \epsilon_{iJ3} \end{bmatrix},$$

with $\boldsymbol{\epsilon_i} \sim N_{3J}(\boldsymbol{0}, \boldsymbol{V})$, and with $\boldsymbol{V} = \boldsymbol{I_J} \otimes \boldsymbol{\Sigma}$,
where $\boldsymbol{\Sigma}$ is the (3×3) covariance matrix of the outcomes,
and with $\boldsymbol{\beta_i} \sim N_6(\boldsymbol{0}, \boldsymbol{\Omega})$.

Assume $\boldsymbol{X}, \boldsymbol{\alpha}, \boldsymbol{W}, \boldsymbol{\Sigma}, \boldsymbol{\Omega}$ are known. To integrate out $\boldsymbol{\beta}$, use the relationship

$$f(\boldsymbol{y}) = \int f(\boldsymbol{y}|\boldsymbol{\beta}) f(\boldsymbol{\beta}) d\boldsymbol{\beta}.$$

The densities are

$$f(\boldsymbol{y}|\boldsymbol{\beta}) \propto \exp\left\{-\tfrac{1}{2} Q_1(\boldsymbol{y}, \boldsymbol{\beta})\right\} \text{ and } f(\boldsymbol{\beta}) \propto \exp\left\{-\tfrac{1}{2} Q_2(\boldsymbol{\beta})\right\},$$

where, dropping subscripts and letting $\boldsymbol{u} = \boldsymbol{y} - \boldsymbol{X\alpha}$, we have

$$Q_1(\boldsymbol{u}, \boldsymbol{\beta}) = (\boldsymbol{u} - \boldsymbol{W\beta})' \boldsymbol{V}^{-1} (\boldsymbol{u} - \boldsymbol{W\beta})$$
$$Q_2(\boldsymbol{\beta}) = \boldsymbol{\beta}' \boldsymbol{\Omega}^{-1} \boldsymbol{\beta}$$

so the joint density is

$$\propto \exp\left\{-\tfrac{1}{2} Q(\boldsymbol{u}, \boldsymbol{\beta})\right\}$$

where $Q(\boldsymbol{u}, \boldsymbol{\beta}) = Q_1(\boldsymbol{u}, \boldsymbol{\beta}) + Q_2(\boldsymbol{\beta})$.

$$\begin{aligned} Q_1(\boldsymbol{u}, \boldsymbol{\beta}) + Q_2(\boldsymbol{\beta}) &= (\boldsymbol{u} - \boldsymbol{W\beta})' \boldsymbol{V}^{-1} (\boldsymbol{u} - \boldsymbol{W\beta}) + \boldsymbol{\beta}' \boldsymbol{\Omega}^{-1} \boldsymbol{\beta} \\ &= u'V^{-1}u - 2u'V^{-1}W\beta + \beta'W'V^{-1}W\beta + \beta'\Omega^{-1}\beta \\ &= \beta'[\Omega^{-1} + W'V^{-1}W]\beta - 2[u'V^{-1}W]\beta + u'V^{-1}u \end{aligned}$$

Letting $A = \Omega^{-1} + W'V^{-1}W$ and $b' = u'V^{-1}W$, we have

$$= (\boldsymbol{\beta} - \boldsymbol{A}^{-1}\boldsymbol{b})' A(\boldsymbol{\beta} - \boldsymbol{A}^{-1}\boldsymbol{b}) - \boldsymbol{b}'\boldsymbol{A}^{-1}\boldsymbol{b} + \boldsymbol{u}'\boldsymbol{V}^{-1}\boldsymbol{u}.$$

Now if we integrate over $\boldsymbol{\beta}$, the kernel

$$\propto \exp\left\{-\frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{A}^{-1}\boldsymbol{b})'\boldsymbol{A}(\boldsymbol{\beta} - \boldsymbol{A}^{-1}\boldsymbol{b})\right\}$$

becomes 1, and we have

$$Q(\boldsymbol{u}) = \boldsymbol{u}'\boldsymbol{V}^{-1}\boldsymbol{u} - \boldsymbol{b}'\boldsymbol{A}^{-1}\boldsymbol{b}$$

$$= \boldsymbol{u}'\boldsymbol{V}^{-1}\boldsymbol{u} - \boldsymbol{u}'\boldsymbol{V}^{-1}\boldsymbol{W}(\boldsymbol{\Omega}^{-1} + \boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{W}\,)^{-1}\boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{u}$$

$$= \boldsymbol{u}'[\boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{W}(\boldsymbol{\Omega}^{-1} + \boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{W}\,)^{-1}\boldsymbol{W}'\boldsymbol{V}^{-1}]\boldsymbol{u}$$

which means that $\boldsymbol{u} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{C}^{-1})$,

and so $\boldsymbol{y} \sim N(\boldsymbol{X\alpha}, \boldsymbol{C}^{-1})$,

where $\boldsymbol{C} = \boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{W}(\boldsymbol{\Omega}^{-1} + \boldsymbol{W}'\boldsymbol{V}^{-1}\boldsymbol{W}\,)^{-1}\boldsymbol{W}'\boldsymbol{V}^{-1}$.