Recall BLUP for \( \hat{u} : \hat{u} = \mathbf{GZ}\mathbf{P}^{-1}\mathbf{y} \)

(and approximation for it \( \hat{u} = \mathbf{GZ}\mathbf{P}^{-1}\mathbf{y} \))

More on the sense in which \( \hat{u} \) is "best" — among all linear predictors it "minimizes" the covariance matrix for \( \mathbf{u} - \hat{\mathbf{u}} \) — an arbitrary linear predictor over choices of \( \hat{\mathbf{u}} \) with mean \( \mu \) — i.e. the BLUP makes small

\[
\text{Var}(\mathbf{u} - \hat{\mathbf{u}})
\]

BTW for the BLUP

\[
\text{Var}(\mathbf{u} - \hat{\mathbf{u}}) = \mathbf{GZPZG}
\]

Both \( \hat{\mathbf{u}} \) and \( \text{Var}(\mathbf{u} - \hat{\mathbf{u}}) \) depend upon \( \Sigma^2 \) and are therefore not realizable/not available for use — but, if I can estimate \( \Sigma^2 \) (via ML or REML) I can get \( \hat{\mathbf{u}} \) and hope that placing hats on \( \mathbf{G} \) and \( \mathbf{P} \) up there will get me something I can use.
to approximate/estimate \( \text{Var}(\bar{u} - \hat{u}) \) i.e.

\[
\text{Var}(\bar{u} - \hat{u}) = \hat{G} - \hat{G}Z'Z \hat{G}
\]

I hope

\[
\text{Var}(\bar{u} - \hat{u})
\]

Further, it is possible to think about prediction of quantities like

\[
l = \frac{e'\varepsilon}{\hat{\alpha}} + \frac{u'\varepsilon}{\hat{\alpha}} > a \text{ linear combination of fixed and random effects}
\]

(this has the obvious generalization to \( l = C\beta + Su \))

This generalizes both estimation of \( \varepsilon'\beta \) and prediction of elements of \( \mu \)

As it turns out, provided \( \varepsilon'\beta \) is estimable \( (\varepsilon' = \alpha'X) \) the BLUP of \( l \) is
\[ \hat{\lambda} = \alpha' \sum \left( \frac{1}{\sigma^2} \right) + \xi' \sum \sim \]

BLUE of EY \[ \sim \]

and a prediction variance for \( \hat{\lambda} \) is

\[ \text{Var}(\hat{\lambda} - \xi) = \xi' (X'V^{-1}X)^{-1} \xi + \xi' G Z' P G \xi - 2 \xi' B Z G \xi \]

Both \( \hat{\lambda} \) and \( \text{Var}(\hat{\lambda} - \xi) \) depend upon unknown \( \sigma^2 \) and are thus not realizable... in the "obvious" way we'll use

\[ \hat{\lambda} = \alpha' \sum \left( \frac{1}{\sigma^2} \right) + \xi' \sum \sim \]

and I'll try to get an empirical approximation to \( \text{Var}(\hat{\lambda} - \xi) \) using

\[ \text{Var}(\hat{\lambda} - \xi) \]

This means plugging \( \sigma^2 \) in for \( \sigma^2 \) in the earlier formula for \( \text{Var}(\hat{\lambda} - \xi) \)

We also need some kind of measure of uncertainty, precision for our estimates of elements of \( \sigma^2 \) — in theory, one could apply
Theory about "The Shape of the Likelihood" (inversion of LRT's) (apply to restricted likelihood) — apparently that is really hard in practice — instead, state of the art software seems to rely on large sample theory for MLEs — (see again Section 7.3.1 of typed outline for a version of this) — recall

\( \hat{\theta} \) some \( r \)-dimensional vector of parameters and

\[ l(\theta) \]

is a log likelihood based on a "large sample" and

\[ \hat{\theta} \] is a maximizer of \( l(\theta) \)

Standard large sample theory then suggests that \( \operatorname{Var}(\hat{\theta}) \) can be approximated/estimated using 2nd derivatives of \( l(\theta) \) at the MLE \( \hat{\theta} \); i.e.

\[
\operatorname{Var}(\hat{\theta}) = \left(-\frac{\partial^2 l(\theta)}{\partial \theta \partial \theta^\prime}\right)^{-1}{_{\theta = \hat{\theta}}}
\]

so, for example, to set approximate confidence limits for \( \theta \).
\( \hat{\theta}_1 \pm z \sqrt{\text{1st diagonal entry of inverse Hessian matrix}} \)