$y = X\beta + Zb + \varepsilon$, where $b \sim N(0, \psi_\theta)$, $\varepsilon \sim N(0, I\sigma^2)$.

$\psi_\theta = \text{diag}(\quad)$ $\Rightarrow$ random effects are independent.

$b_i = b(\theta)$. $\Rightarrow$ set this up

$E(y) = X\beta$ $\quad\quad$ $\text{cov}(y) = Z\psi_\theta Z^T + I\sigma^2$

$y \sim N(X\beta, Z\psi_\theta Z^T + I\sigma^2)$

$$l(\theta, \beta) = \frac{-(y-X\beta)^T(Z\psi_\theta Z^T + I\sigma^2)^{-1}(y-X\beta)}{2} - \frac{\log|Z\psi_\theta Z^T + I\sigma^2|}{2}$$

$\downarrow$ inefficient if $n > p$

$\Rightarrow Z = Q\begin{bmatrix} R \\ 0 \end{bmatrix}$

$$Z\psi_\theta Z^T + I\sigma^2 = Q\begin{bmatrix} R\psi_\theta R^T + I_p\sigma^2 & 0 \\ 0 & I_{n-p}\sigma^2 \end{bmatrix} Q^T.$$

$Q$: first $p$ column of $Q$

$$W = (Z\psi_\theta Z^T + I\sigma^2)^{-1} = Q\begin{bmatrix} (R\psi_\theta R^T + I_p\sigma^2)^{-1} & 0 \\ 0 & I_{np}\sigma^{-2} \end{bmatrix} Q^T$$

$Q^T = Q^{-1}$

$Q^{-T} = Q$

let $R\psi_\theta R^T + I_p\sigma^2 = M^T M$.

$M^{-1}M^{-T}$

$$\log|Z\psi_\theta Z^T + I\sigma^2| = \log|R\psi_\theta R^T + I_p\sigma^2| + (n-p)\log(\sigma^2) = 2\log|S| + 2(n-p)\log\sigma$$

$$\hat\beta = (X^T W X)^{-1} X^T W y$$

let $Q_p$: first $p$ columns of $Q$, $Q_{n-p}$: last $n-p$ columns of $Q$.

$y_p$: first $p$ row of $y$. $y_{n-p}$: last $n-p$ row of $y$

$$X^T W y = \begin{bmatrix} X_p Q_p S^{-1} \underbrace{S^{-T} Q_p^T y_p}_{\text{forwardsolve}} \\ X_{np} Q_{n-p} \cdot Q_{n-p}^T y_{n-p} \cdot \sigma^{-2} \end{bmatrix}$$

$R\psi_\theta R^T + I_p\sigma^2 = S^T S$ $\quad\quad$ $Q^T y = \text{qty}$ $\quad\quad$ $Q^T x = \text{qtx}$

$A = (R\psi_\theta R^T + I_p\sigma^2)^{-1} = S^{-1}S^{-T}$

$$X^T W X = \begin{bmatrix} X_p Q_p S^{-1} \cdot S^{-T} Q_p^T X_{p \times n} \\ X_{np} Q_{n-p} \cdot Q_{n-p}^T X_{(n-p)\times n}\sigma^{-2} \end{bmatrix}.$$

To calculate $\hat\beta$:

let $X^T W X = L^T L$

loglike: $-(y-X\beta)^T Q\begin{bmatrix} A & 0 \\ 0 & I\sigma^{-2} \end{bmatrix} Q^T(y-X\beta)/2 \;=$ change all $X$ into $(y-X\beta)$

$- \log|S| - (n-p)\log\sigma$.

$\therefore \underbrace{L^T L \cdot \hat\beta}_{z} = X^T W y$.

$$\begin{cases} L^T z = X^T W y. \\ L\hat\beta = z. \end{cases}$$