$$\widehat{\beta} = (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}$$

Zeige erst:

$$\begin{aligned}(\boldsymbol{y}-\boldsymbol{X}\beta)'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X}\beta) &= (\boldsymbol{y}-\boldsymbol{X}\widehat{\beta})'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X}\widehat{\beta}) \\ &\quad +(\beta-\widehat{\beta})'(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})(\beta-\widehat{\beta}).\end{aligned} \tag{1}$$

Beweis:

$$\begin{aligned}(\boldsymbol{y}-\boldsymbol{X}\beta)'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X}\beta) &= (\boldsymbol{y}-\boldsymbol{X}\widehat{\beta}+\boldsymbol{X}\widehat{\beta}-\boldsymbol{X}\beta)'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X}\widehat{\beta}+\boldsymbol{X}\widehat{\beta}-\boldsymbol{X}\beta) \\ &= (\boldsymbol{y}-\boldsymbol{X}\widehat{\beta})'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X}\widehat{\beta}) \\ &\quad +(\beta-\widehat{\beta})'(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})(\beta-\widehat{\beta}) \\ &\quad +2(\widehat{\beta}-\beta)'\underbrace{\boldsymbol{X}'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X}\widehat{\beta})}_{\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}-\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}=\boldsymbol{0}}\end{aligned}$$

Herleitung der REML-Schätzung

Matrizen, die die Eigenschaften 1-4 erfüllen sind z.B. $\boldsymbol{A} := \boldsymbol{I_n} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$ und $\boldsymbol{B} := (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}$. Es ist $\boldsymbol{AX} = \boldsymbol{0} \Rightarrow \mathrm{E}(\boldsymbol{Ay}) = \boldsymbol{0}$.

Sei $\boldsymbol{P}$ eine $n \times (n-p)$ Matrix mit $\boldsymbol{PP}' = \boldsymbol{A}$ und $\boldsymbol{P}'\boldsymbol{P} = \boldsymbol{I_{n-p}}$ (Spektralzerlegung einer Projektionsmatrix). Sei $\boldsymbol{z} := \boldsymbol{P}'\boldsymbol{y}$. Gegeben ϑ gilt

$$\widehat{\beta} = (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} = \boldsymbol{By} \sim N(\beta, (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}).$$

Dichten für $\boldsymbol{y}$ und $\widehat{\beta}$:

$$\begin{aligned} p(\boldsymbol{y}) &= (2\pi)^{-n/2}|\boldsymbol{V}|^{-1/2}\exp\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{X}\beta)'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X}\beta)\} \\ p(\widehat{\beta}) &= (2\pi)^{-p/2}|\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}|^{1/2}\exp\{-\frac{1}{2}(\widehat{\beta}-\beta)'(\boldsymbol{X}\boldsymbol{V}^{-1}\boldsymbol{X})(\widehat{\beta}-\beta)\}. \end{aligned}$$

Nun ist

$$\mathrm{E}(\boldsymbol{z}) = \boldsymbol{P}'\boldsymbol{X}\beta = \underbrace{\boldsymbol{P}'\boldsymbol{P}}_{\boldsymbol{I_{n-p}}}\boldsymbol{P}'\boldsymbol{X}\beta = \boldsymbol{P}'\underbrace{\boldsymbol{AX}}_{\boldsymbol{0}}\beta = \boldsymbol{0}.$$

Es folgt

$$\begin{aligned}\operatorname{Cov}(\boldsymbol{z}, \widehat{\boldsymbol{\beta}}) &= \operatorname{Cov}(\boldsymbol{P}'\boldsymbol{y}, \boldsymbol{B}\boldsymbol{y}) = \boldsymbol{P}' \operatorname{Cov}(\boldsymbol{y})\boldsymbol{B}' = \boldsymbol{P}' \underbrace{\boldsymbol{V}\boldsymbol{V}^{-1}}_{\boldsymbol{I}_n} \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1} \\ &= \underbrace{\boldsymbol{P}'\boldsymbol{P}}_{\boldsymbol{I}_{n-p}} \boldsymbol{P}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1} = \boldsymbol{P}' \underbrace{\boldsymbol{A}\boldsymbol{X}}_{\boldsymbol{0}}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1} = \boldsymbol{0}.\end{aligned}$$

Da $(\boldsymbol{z}', \widehat{\boldsymbol{\beta}}')'$ normalverteilt ist, sind $\boldsymbol{z}$ und $\widehat{\boldsymbol{\beta}}$ damit unabhängig. Damit ist die Dichte von $\boldsymbol{z}$ proportional zu

$$\frac{p(\boldsymbol{y})}{p(\widehat{\boldsymbol{\beta}})} \stackrel{(1)}{=} \frac{1}{(2\pi)^{(n-p)/2}|\boldsymbol{V}|^{1/2}|\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}|^{1/2}} \exp\{-\frac{1}{2}(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}})'\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}})\}.$$

Man kann zeigen, dass die Proportionalitätskonstante (Jacobi-Determinante) $|\boldsymbol{X}'\boldsymbol{X}|^{-1/2}$ und damit unabhängig von $\boldsymbol{\vartheta}$ ist.
Diese Likelihood ist unabhängig von der genauen Wahl von $\boldsymbol{A}$.
Beachte, dass Maximierung von $p(\widehat{\boldsymbol{\beta}})$ bzgl. $\boldsymbol{\beta}$ den ML-Schätzer $\widehat{\boldsymbol{\beta}}$ ergibt.