Theorem. If $Y \sim N_n(X\beta, \sigma^2 I_n)$, where $X_{n\times p}$ has rank r and $\hat{\beta} = (X'X)^{-}X'Y$ is a least squares solution of β,
(i) $X\hat{\beta} \sim N_n(X\beta, \sigma^2 P)$,
(ii) $(\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta) \sim \sigma^2\chi^2_r$
(iii) $X\hat{\beta}$ is independent of RSS $= (Y - X\hat{\beta})'(Y - X\hat{\beta})$. and
(iv) RSS$/\sigma^2 \sim \chi^2_{n-r}$ (independent of $X\hat{\beta}$)

Proof. (i) Since $X\hat{\beta} = X(X'X)^{-}X'Y = PY$, we have

$$X\hat{\beta} \sim N_n(PX\beta, \sigma^2 P^2) = N_n(X\beta, \sigma^2 P).$$

(ii) Since $X\hat{\beta} = PY$ and $X\beta = PX\beta$,

$$\begin{aligned}(\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta) &= (X\hat{\beta} - X\beta)'(X\hat{\beta} - X\beta) \\ &= (Y - X\beta)'P(Y - X\beta) \sim \sigma^2\chi^2_r,\end{aligned}$$

P being symmetric idempotent of rank r.
(iii) We have $X\hat{\beta} = PY$, RSS $= (Y - X\hat{\beta})'(Y - X\hat{\beta}) = Y'(I - P)Y$ and $P(I - P) = 0$. Therefore independence of $X\hat{\beta}$ and RSS follows.
(iv) Note again that

$$\text{RSS} = Y'(I - P)Y = (Y - X\beta)'(I - P)(Y - X\beta) \sim \sigma^2\chi^2_{n-r},$$

$I - P$ being a projection matrix of rank $n - r$.

Estimability

Consider the Gauss-Markov model again: $Y = X\beta + \epsilon$, with $E(\epsilon) = 0$ and $Cov(\epsilon) = \sigma^2 I_n$. Now suppose rank of X is $r < p$.

Definition. A linear parametric function $a'\beta$ is said to be estimable if it has a linear unbiased estimate $b'Y$.

Theorem. $a'\beta$ is estimable iff $a \in \mathcal{M}_C(X') = \mathcal{M}_C(X'X)$.

Proof. $a'\beta$ is estimable iff there exists b such that $E(b'Y) = a'\beta$ for all $\beta \in \mathcal{R}^p$. i.e., $b'X\beta = a'\beta$ for all $\beta \in \mathcal{R}^p$. i.e., $b'X = a'$ or $a = X'b$ for some $b \in \mathcal{R}^n$.

Theorem (Gauss-Markov). If $a'\beta$ is estimable, and $\hat{\beta}$ is any least squares solution (i.e., solution of $X'X\beta = X'Y$),
(i) $a'\hat{\beta}$ is unique,
(ii) $a'\hat{\beta}$ is the BLUE of $a'\beta$.

Proof. (i) If $a'\beta$ is estimable, $a'\beta = b'X\beta = b'\theta$ for some $b \in \mathcal{R}^n$. Since $\hat{\theta}$ is the unique projection of Y onto $\mathcal{M}_C(X)$, we note $b'\hat{\theta} = b'X\hat{\beta} = a'\hat{\beta}$ is

unique. i.e., if $\tilde{\beta}$ is any other LS solution, then also $b'X\tilde{\beta} = b'X\hat{\beta} = a'\hat{\beta}$.
(ii) If $d'Y$ is any other linear unbiased estimate of $a'\beta$, then
$E(d'Y) = d'X\beta = d'\theta = a'\beta = b'X\beta = b'\theta$ for all $\beta \in \mathcal{R}^p$.
i.e., $d'\theta = b'\theta$ for all $\theta \in \mathcal{M}_C(X)$.
i.e., $(d-b)'\theta = 0$ for all $\theta \in \mathcal{M}_C(X)$, or $(d-b) \perp \mathcal{M}_C(X)$. Consider $P = P_{\mathcal{M}_C(X)} = X(X'X)^- X'$. Then $P(d-b) = 0$ or $Pd = Pb$. Therefore,

$$\begin{aligned} Var(d'Y) - Var(a'\hat{\beta}) &= Var(d'Y) - Var(b'\hat{\theta}) \\ &= Var(d'Y) - Var(b'PY) = Var(d'Y) - Var(d'PY) \\ &= \sigma^2(d'd - d'Pd) = \sigma^2 d'(I-P)d \geq 0, \end{aligned}$$

with equality iff $(I-P)d = 0$ or $d = Pd = Pb$. i.e., $d'Y = b'PY = b'\hat{\theta} = a'\hat{\beta}$.

Remark. Parametric functions $a'\beta$ are estimable when $a \in \mathcal{M}_C(X') =$ Row space of X.

Example. Consider again the model:
$y_{ij} = \mu + \alpha_i + \tau_j + \epsilon_{ij},\ i = 1,2,3,4;\ j = 1,2.$
Suppose comparing τ_1 and τ_2 is of interest. Since

$$Y = \begin{pmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \\ \vdots \\ y_{41} \\ y_{42} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \tau_1 \\ \tau_2 \end{pmatrix} + \epsilon,$$

$\mu+\alpha_i+\tau_j$ is estimable for all i and j. Therefore, $(\mu+\alpha_i+\tau_1) - (\mu+\alpha_i+\tau_2) = \tau_1 - \tau_2$ is estimable.
$(\mu + \alpha_i + \tau_1) - (\mu + \alpha_j + \tau_1) = \alpha_i - \alpha_j$ is estimable.
What else is estimable, apart from linear combinations of these?

Result. If $a'\beta$ is estimable, and $Y \sim N_n(X\beta, \sigma^2 I_n)$, a $100(1-\alpha)\%$ confidence interval for $a'\beta$ is given by

$$a'\hat{\beta} \pm t_{n-r}(1-\alpha/2)\sqrt{a'(X'X)^- a}\sqrt{\text{RSS}/(n-r)}.$$

Proof. Note that $a'\beta = c'X\beta = c'\theta$ for some c. Therefore, $a'\hat{\beta} = c'\hat{\theta} = c'PY \sim N(a'\beta, \sigma^2 c'Pc)$. Now $c'Pc = c'X(X'X)^- X'c = a'(X'X)^- a$. Therefore,

$$\frac{a'\hat{\beta} - a'\beta}{\sqrt{\sigma^2 a'(X'X)^- a}} \sim N(0,1).$$

Further, since $\mathrm{RSS}/\sigma^2 \sim \chi^2_{n-r}$ independent of $X\hat{\beta}$, and hence of $c'X\hat{\beta} = c'\hat{\theta} = a'\hat{\beta}$,

$$\frac{a'\hat{\beta} - a'\beta}{\sqrt{\sigma^2 a'(X'X)^- a}\sqrt{\mathrm{RSS}/(\sigma^2(n-r))}} \sim t_{n-r}.$$

Hence,

$$P\left(|a'\hat{\beta} - a'\beta| \leq t_{n-r}(1-\alpha/2)\sqrt{a'(X'X)^- a}\sqrt{\frac{\mathrm{RSS}}{n-r}}\right) = 1 - \alpha.$$