Prediction Mean Square Error Calculation based on PMSE Improvement

Deduction of PMSE in Predictive Mean Square Error and Stochastic Regressor Variables by SUBHASH C. NARULA

The response variable and the predictor variables follow a joint (k+ 1)-variate normal distribution with unknown mean vector $\mu^* = [\mu_0, \boldsymbol{\mu}']'$, and unknown covariance matrix $\boldsymbol{\Sigma^*} = \begin{bmatrix} \sigma_{00} & \boldsymbol{\sigma}' \\ \boldsymbol{\sigma} & \boldsymbol{\Sigma} \end{bmatrix}$. Let $\boldsymbol{z_1, z_2, ..., z_n}$ be n independent (k-component vector) observations on the predictor variables, $\boldsymbol{x_i = z_i - \bar{z}}$. Let $\boldsymbol{S^*} = \begin{bmatrix} s_{00} & \boldsymbol{s}' \\ \boldsymbol{s} & \boldsymbol{S} \end{bmatrix}$ be sample covariance matrix, where

$$s_{00} = \sum \frac{(y_i - \bar{y})^2}{n-1}, \boldsymbol{s} = \sum \frac{(y_i - \bar{y})\boldsymbol{x_i}}{n-1}, \boldsymbol{S} = \sum \frac{\boldsymbol{x_i x_i'}}{n-1}$$

We assume the correct the model(1.1)

$$\boldsymbol{y = \alpha + \beta_1 z_1 + \beta_2 z_2 + ... + \beta_k z_k + \epsilon}$$

The LSE prediction equation

$$\boldsymbol{\hat{y}} = \bar{y} + \boldsymbol{\hat{\beta}_1}(\boldsymbol{z_1} - \bar{z}_1) + \boldsymbol{\hat{\beta}_2}(\boldsymbol{z_2} - \bar{z}_2) + ... + \boldsymbol{\hat{\beta}_k}(\boldsymbol{z_k} - \bar{z}_k) = \bar{y} + \boldsymbol{X'\hat{\beta}}$$

$$\hat{y}_i = \bar{y} + \boldsymbol{x_i'\hat{\beta}}$$

For any given $\boldsymbol{z_i}$,

$$\begin{aligned} E(y_i|\boldsymbol{z_i}) &= \alpha + \boldsymbol{\beta z_i} \\ &= \mu_0 - \boldsymbol{\sigma}'\boldsymbol{\Sigma^{-1}\mu} + \boldsymbol{\sigma'\Sigma^{-1}z_i} \\ &= \mu_0 + \boldsymbol{\sigma}'\boldsymbol{\Sigma^{-1}(z_i - \mu)} \end{aligned}$$

where $\alpha = \mu_0 - \boldsymbol{\sigma}'\boldsymbol{\Sigma^{-1}\mu}, \boldsymbol{\beta = \sigma}'\boldsymbol{\Sigma^{-1}}$. Thus $\hat{\alpha} = \bar{y} - \boldsymbol{s'S^{-1}\bar{x}}, \boldsymbol{\hat{\beta} = S^{-1}s}$.

The conditional predictive mean square error by

$$
\begin{aligned}
E[(y_0-\hat{y_0})^2|\boldsymbol{z_0}] &= E[(\alpha+(\boldsymbol{z_0}-\boldsymbol{\mu})'\boldsymbol{\beta}+\epsilon_0-\bar{y}-\boldsymbol{x_0'}\hat{\boldsymbol{\beta}}|\boldsymbol{z_0})^2] \\
&= E[(\alpha+(\boldsymbol{z_0}-\boldsymbol{\mu})'\boldsymbol{\beta}+\boldsymbol{\epsilon_0}-\alpha-(\bar{\boldsymbol{z}}-\boldsymbol{\mu})'\boldsymbol{\beta}-\bar{\boldsymbol{\epsilon}}-\boldsymbol{x_0'}\beta|\boldsymbol{z_0})^2] \\
&= E[(\boldsymbol{x_0'}\boldsymbol{\beta}-\bar{\boldsymbol{x}}'\boldsymbol{\beta}+\epsilon_0-\bar{\epsilon}-\boldsymbol{x_0'}\hat{\boldsymbol{\beta}}|\boldsymbol{z_0})^2] \\
&= E[(\boldsymbol{x_0'}\boldsymbol{\beta}+(\epsilon_0-\bar{\epsilon})-\boldsymbol{x_0'}\hat{\boldsymbol{\beta}}|\boldsymbol{z_0})^2] \\
&= E[(\boldsymbol{x_0'}\boldsymbol{\beta}+(\epsilon_0-\bar{\epsilon}))^2+(\boldsymbol{x_0'}\hat{\boldsymbol{\beta}})^2-2(\boldsymbol{x_0'}\boldsymbol{\beta}+(\epsilon_0-\bar{\epsilon}))\boldsymbol{x_0'}\hat{\boldsymbol{\beta}}|\boldsymbol{z_0}] \\
&= E[(\boldsymbol{x_0'}\boldsymbol{\beta})^2+(\epsilon_0-\bar{\epsilon})^2+(\boldsymbol{x_0'}\hat{\boldsymbol{\beta}})^2-2\boldsymbol{x_0'}\boldsymbol{\beta}\boldsymbol{x_0'}\hat{\boldsymbol{\beta}}|\boldsymbol{z_0}] \\
&= \boldsymbol{\beta'}E(\boldsymbol{x_0}\boldsymbol{x_0'}|\boldsymbol{z_0})\boldsymbol{\beta}+E[(\epsilon_0-\bar{\epsilon})^2|\boldsymbol{z_0}]+E[(\boldsymbol{x_0'}\hat{\boldsymbol{\beta}})^2|\boldsymbol{z_0}]-2\boldsymbol{\beta'}E[\boldsymbol{x_0}\boldsymbol{x_0'}\hat{\boldsymbol{\beta}}|\boldsymbol{z_0}]
\end{aligned}
$$

By Lemma A1, Lemma A3, Lemma A7

$$
\begin{aligned}
&E(\tilde{\boldsymbol{\beta}}_{\boldsymbol{1}}|\boldsymbol{X_1}) = \boldsymbol{\beta_1}+\boldsymbol{\Sigma_{11}^{-1}\Sigma_{12}\beta_2} = \boldsymbol{\Phi_1} \\
&E(\boldsymbol{x_{01}}\boldsymbol{x_{01}'}|\boldsymbol{z_0}) = (\boldsymbol{z_{01}}-\boldsymbol{\mu_1})(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})' + \frac{\boldsymbol{\Sigma_{11}}}{n} \\
&E(\boldsymbol{x_0}\boldsymbol{x_0'}|\boldsymbol{z_0}) = (\boldsymbol{z_0}-\boldsymbol{\mu})(\boldsymbol{z_0}-\boldsymbol{\mu})' + \frac{\boldsymbol{\Sigma}}{n} \\
&E[(\boldsymbol{x_{01}'}\tilde{\boldsymbol{\beta}}_{\boldsymbol{1}})^2|\boldsymbol{z_0}] = \sigma_p^2\left[(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})'\boldsymbol{\Sigma_{11}^{-1}}(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})+\frac{p}{n}\right]\frac{1}{n-p-2} \\
&\qquad\qquad + \boldsymbol{\Phi_1'\Sigma_{11}\Phi_1}\frac{1}{n}+\boldsymbol{\Phi_1'}(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})'\boldsymbol{\Phi_1} \\
&\boldsymbol{\sigma'\Sigma^{-1}}\begin{bmatrix}\boldsymbol{\Sigma_{11}}\\ \boldsymbol{\Sigma_{21}}\end{bmatrix} = \sigma_1'
\end{aligned}
$$

The conditional PMSE can be written as

$$
\begin{aligned}
E[(y_0-\hat{y_0})^2|\boldsymbol{z_0}] &= \boldsymbol{\beta'}E(\boldsymbol{x_0}\boldsymbol{x_0'}|\boldsymbol{z_0})\boldsymbol{\beta}+E[(\epsilon_0-\bar{\epsilon})^2|\boldsymbol{z_0}]+E[(\boldsymbol{x_0'}\hat{\boldsymbol{\beta}})^2|\boldsymbol{z_0}]-2\boldsymbol{\beta'}E[\boldsymbol{x_0}\boldsymbol{x_0'}\hat{\boldsymbol{\beta}}|\boldsymbol{z_0}] \\
&= \boldsymbol{\beta}'\left[(\boldsymbol{z_0}-\boldsymbol{\mu})(\boldsymbol{z_0}-\boldsymbol{\mu})'+\frac{\boldsymbol{\Sigma}}{n}\right]\boldsymbol{\beta}+\sigma_k^2\left(1+\frac{1}{n}\right) \\
&\quad + \sigma_k^2\left[(\boldsymbol{z_0}-\boldsymbol{\mu})'\boldsymbol{\Sigma^{-1}}(\boldsymbol{z_0}-\boldsymbol{\mu})+\frac{k}{n}\right]\frac{1}{n-k-2} \\
&\quad + \frac{1}{n}\boldsymbol{\Phi'\Sigma\Phi}+\boldsymbol{\Phi'}(\boldsymbol{z_0}-\boldsymbol{\mu})(\boldsymbol{z_0}-\boldsymbol{\mu})'\boldsymbol{\Phi}-2\boldsymbol{\beta}'\left[(\boldsymbol{z_0}-\boldsymbol{\mu})(\boldsymbol{z_0}-\boldsymbol{\mu})'+\frac{\boldsymbol{\Sigma}}{n}\right]\boldsymbol{\beta} \\
&= \sigma_k^2\left(1+\frac{1}{n}\right)+\sigma_k^2\left[(\boldsymbol{z_0}-\boldsymbol{\mu})'\boldsymbol{\Sigma^{-1}}(\boldsymbol{z_0}-\boldsymbol{\mu})+\frac{k}{n}\right]\frac{1}{n-k-2}
\end{aligned}
$$

Since $\boldsymbol{\Phi}$ is the notation of expectation of $\tilde{\boldsymbol{\beta}}_{\boldsymbol{1}}$, when we are using all predictors, the LSE is unbiased, which means $\boldsymbol{\Phi}=\boldsymbol{\beta}$.

The unconditional PMSE

$$
\begin{aligned}
E[(y_0-\hat{y_0})^2] &= E\{E[(y_0-\hat{y_0})^2|\boldsymbol{z_0}]\} \\
&= E\left[\sigma_k^2\left(1+\frac{1}{n}\right)+\sigma_k^2\left[\boldsymbol{(z_0-\mu)'\Sigma^{-1}(z_0-\mu)}+\frac{k}{n}\right]\frac{1}{n-k-2}\right] \\
&= \sigma_k^2\left(1+\frac{1}{n}\right)+\sigma_k^2 E\left[\boldsymbol{(z_0-\mu)'\Sigma^{-1}(z_0-\mu)}+\frac{k}{n}\right]\frac{1}{n-k-2} \\
&= \sigma_k^2\left(1+\frac{1}{n}\right)+\sigma_k^2\left(k+\frac{k}{n}\right)\frac{1}{n-k-2} \\
&= \sigma_k^2\left(1+\frac{1}{n}\right)\left(1+\frac{1}{n-k-2}\right) \\
&= \sigma_k^2\left(1+\frac{1}{n}\right)\left(\frac{n-2}{n-k-2}\right)
\end{aligned}
$$

For subset, we partition the k-component vector of predictor variables into two parts, $\boldsymbol{Z}=[\boldsymbol{Z_1},\boldsymbol{Z_2}],\boldsymbol{X}=[\boldsymbol{X_1},\boldsymbol{X_2}],\boldsymbol{x_1'}=[\boldsymbol{x_{i1}'},\boldsymbol{x_{i2}'}],\boldsymbol{\mu'}=[\boldsymbol{\mu_1'},\boldsymbol{\mu_2'}],\boldsymbol{\sigma'}=[\boldsymbol{\sigma_1'},\boldsymbol{\sigma_2'}],\boldsymbol{s'}=[\boldsymbol{s_1'},\boldsymbol{s_2'}]$, $\boldsymbol{\Sigma}=\begin{bmatrix}\boldsymbol{\Sigma_{11}} & \boldsymbol{\Sigma_{12}}\\ \boldsymbol{\Sigma_{21}} & \boldsymbol{\Sigma_{22}}\end{bmatrix},\boldsymbol{S}=\begin{bmatrix}\boldsymbol{S_{11}} & \boldsymbol{S_{12}}\\ \boldsymbol{S_{21}} & \boldsymbol{S_{22}}\end{bmatrix}$

, so the subset prediction equation is given by

$$\tilde{y_i}=\bar{y}+\boldsymbol{x_{i1}'\tilde{\beta_1}}$$

where $\boldsymbol{\tilde{\beta_1}=S_{11}^{-1}s_1}$. By LemmaA1, LemmaA3, LemmaA7, $\boldsymbol{\Phi_1=\beta_1+\Sigma_{11}^{-1}\Sigma_{12}\beta_2}$, the conditional PMSE at z_0 is given by

$$
\begin{aligned}
E[(y_0-\tilde{y_0})^2|\boldsymbol{z_0}] &= E[(\boldsymbol{x_0'\beta}+\epsilon_0-\bar{\epsilon}-\boldsymbol{x_{01}'\tilde{\beta_1}}|\boldsymbol{z_0})^2] \\
&= E[(\boldsymbol{x_0'\beta})^2+(\boldsymbol{x_{01}'\tilde{\beta_1}})^2-2\boldsymbol{x_0'\beta x_{01}'\tilde{\beta_1}}+(\epsilon_0-\bar{\epsilon})^2+2(\boldsymbol{x_0'\beta-x_{01}'\tilde{\beta_1}})(\epsilon_0-\bar{\epsilon})|\boldsymbol{z_0}] \\
&= E[(\boldsymbol{x_0'\beta})^2|\boldsymbol{z_0}]+E[\boldsymbol{x_{01}'\tilde{\beta_1}})^2|\boldsymbol{z_0}]-2E[\boldsymbol{x_0'\beta x_{01}'\tilde{\beta_1}}|\boldsymbol{z_0}]+E[(\epsilon_0-\bar{\epsilon})^2|\boldsymbol{z_0}] \\
&\quad +2E[(\boldsymbol{x_0'\beta-x_{01}'\tilde{\beta_1}})(\epsilon_0-\bar{\epsilon})|\boldsymbol{z_0}]
\end{aligned}
$$

where

$$E[(\boldsymbol{x_0'\beta})^2|\boldsymbol{z_0}]=\boldsymbol{\beta}\left[\boldsymbol{(z_0-\mu)(z_0-\mu)'}+\frac{\boldsymbol{\Sigma}}{n}\right]\boldsymbol{\beta}$$

$$
\begin{aligned}
E[(\boldsymbol{x_{01}'\tilde{\beta_1}})^2|\boldsymbol{z_0}] =&\sigma_p^2\left[\boldsymbol{(z_{01}-\mu_1)'\Sigma_{11}^{-1}(z_{01}-\mu_1)}+\frac{p}{n}\right]\frac{1}{n-p-2} \\
&+\boldsymbol{\Phi_1'\Sigma_{11}\Phi_1}\frac{1}{n}+\boldsymbol{\Phi_1'(z_{01}-\mu_1)(z_{01}-\mu_1)'\Phi_1},
\end{aligned}
$$

$$E[(\epsilon_0-\bar{\epsilon})^2|\boldsymbol{z_0}]=\sigma_k^2+\frac{1}{n}\sigma_k^2,$$

$$E[(\boldsymbol{x_0'\beta-x_{01}'\tilde{\beta_1}})(\epsilon_0-\bar{\epsilon})|\boldsymbol{z_0}]=E[(\boldsymbol{x_0'\beta-x_{01}'\tilde{\beta_1}})\epsilon_0|\boldsymbol{z_0}]-E[(\boldsymbol{x_0'\beta-x_{01}'\tilde{\beta_1}})\bar{\epsilon}|\boldsymbol{z_0}]=0$$

(but $\bar{\epsilon}=0$ is not necessary)

Equation 3.6 could be written as

$$
\begin{aligned}
&= \sigma_p^2\left[(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})'\boldsymbol{\Sigma_{11}^{-1}}(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})+\frac{p}{n}\right]\frac{1}{n-p-2}+\boldsymbol{\Phi_1'\Sigma_{11}\Phi_1}\frac{1}{n}+\boldsymbol{\Phi_1'}(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})'\boldsymbol{\Phi_1}\\
&\quad+\boldsymbol{\beta}(\boldsymbol{z_0}-\boldsymbol{\mu})(\boldsymbol{z_0}-\boldsymbol{\mu})'\boldsymbol{\beta}+\boldsymbol{\beta'\Sigma\beta}\frac{1}{n}-2\boldsymbol{\beta'}E(\boldsymbol{x_0x_{01}'}|\boldsymbol{z_0})\boldsymbol{\Phi_1}+\sigma_k^2+\frac{1}{n}\sigma_k^2\\
&= \sigma_k^2+\frac{1}{n}(\sigma_p^2+\boldsymbol{\sigma_1'\Sigma_{11}^{-1}\sigma_1}-\boldsymbol{\sigma'\Sigma^{-1}\sigma})+[(\boldsymbol{z_0}-\boldsymbol{\mu})'\beta]^2+[(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})'\boldsymbol{\Phi_1}]^2\\
&\quad+\sigma_p^2\left[(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})'\boldsymbol{\Sigma_{11}^{-1}}(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})+\frac{p}{n}\right]\frac{1}{n-p-2}\\
&\quad+\boldsymbol{\beta'\Sigma\beta}\frac{1}{n}+\boldsymbol{\Phi_1'\Sigma_{11}\Phi_1}\frac{1}{n}-2\boldsymbol{\beta'}E(\boldsymbol{x_0x_{01}'}|\boldsymbol{z_0})\boldsymbol{\Phi_1}\\
&= \sigma_k^2+\frac{1}{n}\sigma_p^2+\sigma_p^2\left[(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})'\boldsymbol{\Sigma_{11}^{-1}}(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})+\frac{p}{n}\right]\frac{1}{n-p-2}+[(\boldsymbol{z_0}-\boldsymbol{\mu})'\beta]^2+[(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})'\boldsymbol{\Phi_1}]^2\\
&\quad+\frac{1}{n}\boldsymbol{\beta'\Sigma\beta}+\frac{1}{n}\boldsymbol{\Phi_1'\Sigma_{11}\Phi_1}+\frac{1}{n}\boldsymbol{\sigma_1'\Sigma_{11}^{-1}\sigma_1}-\frac{1}{n}\boldsymbol{\sigma'\Sigma^{-1}\sigma}-2\boldsymbol{\beta'}E(\boldsymbol{x_0x_{01}'}|\boldsymbol{z_0})\boldsymbol{\Phi_1}
\end{aligned}
$$

where

$$
E(\boldsymbol{x_0x_{01}'}|\boldsymbol{z_0}) = E\left[\begin{bmatrix}\boldsymbol{x_{01}x_{01}'}\\ \boldsymbol{x_{02}x_{01}'}\end{bmatrix}|\boldsymbol{z_0}\right] = \begin{bmatrix}(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})'+\boldsymbol{\Sigma_{11}/n}\\ (\boldsymbol{z_{02}}-\boldsymbol{\mu_2})(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})'+\boldsymbol{\Sigma_{21}/n}\end{bmatrix}
$$

$$
\boldsymbol{\sigma'\Sigma^{-1}\sigma}=\boldsymbol{\beta'}\boldsymbol{\Sigma}\beta,\boldsymbol{\sigma_1'\Sigma_{11}^{-1}\sigma_1}=\beta_1'\boldsymbol{\Sigma_{11}\beta_1}
$$

so

$$
\begin{aligned}
\boldsymbol{\beta'}E(\boldsymbol{x_0x_{01}'}|\boldsymbol{z_0})\boldsymbol{\Phi_1} =&[\boldsymbol{\beta_1'}(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})'+\beta_1'\boldsymbol{\Sigma_{11}}/n+\beta_2'(\boldsymbol{z_{02}}-\boldsymbol{\mu_2})(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})'+\beta_2'\boldsymbol{\Sigma_{21}}/n]\boldsymbol{\Phi_1}\\
=&\boldsymbol{\beta_1'}(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})'\boldsymbol{\Phi_1}+\beta_2'(\boldsymbol{z_{02}}-\boldsymbol{\mu_2})(\boldsymbol{z_{01}}-\boldsymbol{\mu_1})'\boldsymbol{\Phi_1}+\boldsymbol{\sigma_1'\Phi_1}/n
\end{aligned}
$$

In addition to a few terms appearing in 3.6a, other terms would be equal to

$$
\begin{aligned}
&\frac{1}{n}\boldsymbol{\beta'\Sigma\beta}+\frac{1}{n}\boldsymbol{\Phi_1'\Sigma_{11}\Phi_1}+\frac{1}{n}\boldsymbol{\sigma_1'\Sigma_{11}^{-1}\sigma_1}-\frac{1}{n}\boldsymbol{\sigma'\Sigma^{-1}\sigma}-\frac{2}{n}\boldsymbol{\sigma_1'\Phi_1}\\
=&\frac{1}{n}\boldsymbol{\Phi_1'\Sigma_{11}\Phi_1}+\frac{1}{n}\boldsymbol{\sigma_1'\Sigma_{11}^{-1}\sigma_1}-\frac{2}{n}\boldsymbol{\sigma_1'\Phi_1}\\
=&\frac{1}{n}\boldsymbol{\sigma_1'\Sigma_{11}^{-1}\sigma_1}+\frac{1}{n}\boldsymbol{\sigma_1'\Sigma_{11}^{-1}\sigma_1}-\frac{2}{n}\boldsymbol{\sigma_1'\Sigma_{11}^{-1}\sigma_1}=0
\end{aligned}
$$

in which $\Phi_1=\boldsymbol{\Sigma_{11}^{-1}\sigma_1}$

The conditional PMSE is equal to (3.6a)

$$
\begin{aligned}
E[(y_0-\tilde{y_0})^2|\boldsymbol{z_0}] &= \sigma_k^2+\frac{\sigma_p^2}{n}+\sigma_p^2\left[\boldsymbol{(z_{01}-\mu_1)'\Sigma_{11}^{-1}(z_{01}-\mu_1)'}+\frac{p}{n}\right]\frac{1}{n-p-2}\\
&\quad+[\boldsymbol{(z_0-\mu)'\beta-(z_{01}-\mu_1)'\Phi_1}]^2\\
&= \sigma_k^2+\frac{\sigma_p^2}{n}+\sigma_p^2\left[\boldsymbol{(z_{01}-\mu_1)'\Sigma_{11}^{-1}(z_{01}-\mu_1)'}+\frac{p}{n}\right]\frac{1}{n-p-2}\\
&\quad+[\boldsymbol{(z_{01}-\mu_1)'\beta_1+(z_{02}-\mu_2)'-(z_{01}-\mu_1)'(\beta_1+\Sigma_{11}^{-1}\Sigma_{12}\beta_2)}]^2\\
&= \sigma_k^2+\frac{\sigma_p^2}{n}+\sigma_p^2\left[\boldsymbol{(z_{01}-\mu_1)'\Sigma_{11}^{-1}(z_{01}-\mu_1)'}+\frac{p}{n}\right]\frac{1}{n-p-2}\\
&\quad+[\boldsymbol{(z_{02}-\mu_2)'-(z_{01}-\mu_1)'\Sigma_{11}^{-1}\Sigma_{12}\beta_2}]^2
\end{aligned}
$$

Take expectation

$$
\begin{aligned}
E[(y_0-\tilde{y_0})^2] &= E[E(y_0-\tilde{y_0})^2|\boldsymbol{z_0}]\\
&= E\{\sigma_k^2+\frac{\sigma_p^2}{n}+\sigma_p^2\left[\boldsymbol{(z_{01}-\mu_1)'\Sigma_{11}^{-1}(z_{01}-\mu_1)'}+\frac{p}{n}\right]\frac{1}{n-p-2}\\
&\quad+[\boldsymbol{(z_0-\mu)'\beta-(z_{01}-\mu_1)'\Phi_1}]^2\}\\
&= \sigma_p^2+\frac{\sigma_p^2}{n}+\boldsymbol{\sigma_1'\Sigma_{11}^{-1}\sigma_1-\sigma'\Sigma^{-1}\sigma}+\sigma_p^2\left(p+\frac{p}{n}\right)\frac{1}{n-p-2}\\
&\quad+E[\boldsymbol{\beta'(z_0-\mu)(z_0-\mu)'\beta+\Phi_1'(z_{01}-\mu_1)(z_{01}-\mu_1)'\Phi_1}\\
&\quad-2\boldsymbol{\beta'(z_0-\mu)(z_{01}-\mu_1)'\Phi_1}]
\end{aligned}
$$

The expectation term is equal to

$$
\boldsymbol{\beta'\Sigma\beta+\Phi_1'\Sigma_{11}\Phi_1-2\beta'}\begin{bmatrix}\boldsymbol{\Sigma_{11}}\\ \boldsymbol{\Sigma_{21}}\end{bmatrix}\boldsymbol{\Phi_1}
$$

The unconditional PMSE

$$
\begin{aligned}
E[(y_0-\tilde{y_0})^2] &= \sigma_p^2(1+\frac{1}{n})(n-2)/(n-p-2)\\
&\quad+\boldsymbol{\sigma_1'\Sigma_{11}^{-1}\sigma_1-\sigma'\Sigma^{-1}\sigma}+\beta'\boldsymbol{\Sigma}\beta+\boldsymbol{\Phi_1'\Sigma_{11}\Phi_1-2\beta'}\begin{bmatrix}\boldsymbol{\Sigma_{11}}\\ \boldsymbol{\Sigma_{21}}\end{bmatrix}\boldsymbol{\Phi_1}\\
&= \sigma_p^2(1+\frac{1}{n})(n-2)/(n-p-2)\\
&\quad+\boldsymbol{\Phi_1'\Sigma_{11}\Phi_1-\beta'\Sigma\beta}+\beta'\boldsymbol{\Sigma}\beta+\boldsymbol{\Phi_1'\Sigma_{11}\Phi_1-2\beta'}\begin{bmatrix}\boldsymbol{\Sigma_{11}}\\ \boldsymbol{\Sigma_{21}}\end{bmatrix}\boldsymbol{\Phi_1}\\
&= \sigma_p^2(1+\frac{1}{n})(n-2)/(n-p-2)
\end{aligned}
$$

Thus, the unconditional PMSE $= \sigma_p^2(1 + \frac{1}{n})(n-2)/(n-p-2)$