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 $\boldsymbol{\Sigma}^* = \begin{bmatrix} \sigma_{00} & \boldsymbol{\sigma}' \\ \boldsymbol{\sigma} & \Sigma \end{bmatrix}$. 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

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

For any given $\boldsymbol{z_i}$,

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

wrong notation, not a scalar

different, since x_i = z_i - \bar{z}?

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

The LSE prediction is given by

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

and the conditional predictive mean square error by

$$\begin{aligned} E[(y_0 - \hat{y}_0)^2|\boldsymbol{z_0}] &= E[(\alpha + \boldsymbol{x_0'\beta} + \epsilon_0 - \bar{y} - \boldsymbol{x_0'\hat{\beta}}|\boldsymbol{z_0})^2] \\ &= E[(\alpha + \boldsymbol{x_0'\beta} + \epsilon_0 - \alpha + \boldsymbol{\bar{x}'\beta} - \bar{\epsilon} - \boldsymbol{x_0'\hat{\beta}}|\boldsymbol{z_0})^2] \\ &= E[(\boldsymbol{x_0'\beta} + \boldsymbol{\bar{x}'\beta} + \epsilon_0 - \bar{\epsilon} - \boldsymbol{x_0'\hat{\beta}}|\boldsymbol{z_0})^2] \\ &= E[(\boldsymbol{x_0'\beta} + (\epsilon_0 - \bar{\epsilon}) - \boldsymbol{x_0'\hat{\beta}}|\boldsymbol{z_0})^2] \\ &= E[(\boldsymbol{x_0'\beta} + (\epsilon_0 - \bar{\epsilon}))^2 + (\boldsymbol{x_0'\hat{\beta}})^2 - 2(\boldsymbol{x_0'\beta} + (\epsilon_0 - \bar{\epsilon}))\boldsymbol{x_0'\hat{\beta}}|\boldsymbol{z_0}] \\ &= E[(\boldsymbol{x_0'\beta})^2 + (\epsilon_0 - \bar{\epsilon})^2 + (\boldsymbol{x_0'\hat{\beta}})^2 - 2\boldsymbol{x_0'\beta x_0'\hat{\beta}}|\boldsymbol{z_0}] \\ &= \boldsymbol{\beta}'E(\boldsymbol{x_0x_0'}|\boldsymbol{z_0})\boldsymbol{\beta} + E[(\epsilon_0 - \bar{\epsilon})^2|\boldsymbol{z_0}] + E[(\boldsymbol{x_0'\hat{\beta}})^2|\boldsymbol{z_0}] - 2\boldsymbol{\beta}'E[\boldsymbol{x_0x_0'\hat{\beta}}|\boldsymbol{z_0}] \end{aligned}$$

\bar{y} contains \bar{z} instead of \bar{x}=0?

By Lemma A1, Lemma A3, Lemma A7

$$\begin{aligned} E(\boldsymbol{\tilde{\beta}_1}|\boldsymbol{X_1}) &= \boldsymbol{\beta_1} + \boldsymbol{\Sigma_{11}^{-1}\Sigma_{12}\beta_2} = \boldsymbol{\Phi_1} \\ E(\boldsymbol{x_{01}x_{01}'}|\boldsymbol{z_0}) &= (\boldsymbol{z_{01}} - \boldsymbol{\mu_1})(\boldsymbol{z_{01}} - \boldsymbol{\mu_1})' + \frac{\boldsymbol{\Sigma_{11}}}{n} \\ E(\boldsymbol{x_0x_0'}|\boldsymbol{z_0}) &= (\boldsymbol{z_0 - \mu})(\boldsymbol{z_0 - \mu})' + \frac{\boldsymbol{\Sigma}}{n} \\ E[(\boldsymbol{x_{01}'\tilde{\beta}_1})^2|\boldsymbol{z_0}] &= \sigma_p^2 \left[(\boldsymbol{z_{01} - \mu_1})'\boldsymbol{\Sigma_{11}^{-1}}(\boldsymbol{z_{01} - \mu_1}) + \frac{p}{n} \right] \frac{1}{n-p-2} \\ &\quad + \boldsymbol{\Phi_1'\Sigma_{11}\Phi_1}\frac{1}{n} + \boldsymbol{\Phi_1'}(\boldsymbol{z_{01} - \mu_1})(\boldsymbol{z_{01} - \mu_1})'\boldsymbol{\Phi_1} \\ \boldsymbol{\sigma}'\boldsymbol{\Sigma}^{-1} \begin{bmatrix} \boldsymbol{\Sigma_{11}} \\ \boldsymbol{\Sigma_{21}} \end{bmatrix} &= \boldsymbol{\sigma_1'} \end{aligned}$$

The conditional PMSE can be written as

$$\begin{aligned}
E[(y_0-\hat{y_0})^2|\boldsymbol{z_0}] &= \boldsymbol{\beta' } E(\boldsymbol{x_0x_0'}|\boldsymbol{z_0})\boldsymbol{\beta}+E[(\epsilon_0-\bar{\epsilon})^2|\boldsymbol{z_0}]+E[(\boldsymbol{x_0'\hat{\beta}})^2|\boldsymbol{z_0}]-2\boldsymbol{\beta'}E[\boldsymbol{x_0x_0'\hat{\beta}}|\boldsymbol{z_0}]\\
&=\boldsymbol{\beta'}\left[(\boldsymbol{z_0}-\boldsymbol{\mu})(\boldsymbol{z_0}-\boldsymbol{\mu})'+\frac{\boldsymbol{\Sigma}}{n}\right]\boldsymbol{\beta}+\sigma_k^2\left(1+\frac{1}{n}\right)\\
&\quad+\sigma_k^2\left[(\boldsymbol{z_0}-\boldsymbol{\mu})'\boldsymbol{\Sigma^{-1}}(\boldsymbol{z_0}-\boldsymbol{\mu})+\frac{k}{n}\right]\frac{1}{n-k-2}\\
&\quad+\frac{1}{n}\boldsymbol{\Phi'\Sigma\Phi}+\boldsymbol{\Phi'}(\boldsymbol{z_0}-\boldsymbol{\mu})(\boldsymbol{z_0}-\boldsymbol{\mu})'\boldsymbol{\Phi}-2\boldsymbol{\beta'}\left[(\boldsymbol{z_0}-\boldsymbol{\mu})(\boldsymbol{z_0}-\boldsymbol{\mu})'+\frac{\boldsymbol{\Sigma}}{n}\right]\boldsymbol{\beta}\\
&=\sigma_k^2\left(1+\frac{1}{n}\right)+\sigma_k^2\left[(\boldsymbol{z_0}-\boldsymbol{\mu})'\boldsymbol{\Sigma^{-1}}(\boldsymbol{z_0}-\boldsymbol{\mu})+\frac{k}{n}\right]\frac{1}{n-k-2}
\end{aligned}$$

Since Φ is the notation of expactation of $\tilde{\beta}$, when we are using all preditors, the LSE is unbiased, which means $\Phi=\beta$.

The unconditional PMSE

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

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

\bar{y} contains \bar{z} instead of \bar{x}=0?

$$
\begin{aligned}
E[(y_0 - \tilde{y_0})^2|\boldsymbol{z_0}] &= E[(\boldsymbol{x_0'\beta} + \epsilon_0 - \bar{\epsilon} - \boldsymbol{x_{01}'\tilde{\beta}_1}|\boldsymbol{z_0})^2] \\
&= E[(\epsilon_0 - \bar{\epsilon})^2 + (\boldsymbol{x_0'\beta} - \boldsymbol{x_{01}'\tilde{\beta}_1})^2|\boldsymbol{z_0}] \\
&= E[(\epsilon_0 - \bar{\epsilon})^2|\boldsymbol{z_0}] + Var[(\boldsymbol{x_0'\beta} - \boldsymbol{x_{01}'\tilde{\beta}_1})|\boldsymbol{z_0}] + \{E[(\boldsymbol{x_0'\beta} - \boldsymbol{x_{01}'\tilde{\beta}_1})|\boldsymbol{z_0}]\}^2 \\
&= \sigma_k^2 + \frac{\sigma_p^2}{n} + Var(\boldsymbol{x_0'\beta}|\boldsymbol{z_0}) + Var(\boldsymbol{x_{01}'\tilde{\beta}_1}|\boldsymbol{z_0}) + [(\boldsymbol{z_0} - \boldsymbol{\mu})'\boldsymbol{\beta} - (\boldsymbol{z_{01}} - \boldsymbol{\mu_1})'\boldsymbol{\Phi_1}]^2 \\
&= \sigma_k^2 + \frac{\sigma_p^2}{n} + \sigma_k^2 + \sigma_p^2\left[(\boldsymbol{z_{01}} - \boldsymbol{\mu_1})'\boldsymbol{\Sigma_{11}^{-1}}(\boldsymbol{z_{01}} - \boldsymbol{\mu_1}) + \frac{p}{n}\right]\frac{1}{n-p-2} + \boldsymbol{\Phi_1'\Sigma_{11}\Phi_1}/n \\
&\quad + [(\boldsymbol{z_0} - \boldsymbol{\mu})'\boldsymbol{\beta} - (\boldsymbol{z_{01}} - \boldsymbol{\mu_1})\boldsymbol{\Phi_1}]^2
\end{aligned}
$$

what happened to these two terms?

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