多元线性回归中$var(\hat{y})$随回归元x_0的变化规律

在给定某X_0下，对$y_0\equiv \mathrm{X}_0\beta+\epsilon_0$进行预测：

$$\hat{y_0}=y_0+\mathrm{X}_0(\hat{\beta}-\beta)-\epsilon_0$$
$$var(\hat{y_0})=\mathrm{X}_0 var(\hat{\beta}){\mathrm{X}_0}^{\mathrm{T}}+var(\epsilon_0)=\sigma^2(1+\mathrm{X}_0(\mathrm{X}^{\mathrm{T}}\mathrm{X})^{-1}{\mathrm{X}_0}^{\mathrm{T}})$$

接下来我们要证明

$$\underset{\mathrm{X}_0}{\arg\min}\quad var(\hat{y_0})=\bar{\mathrm{X}}$$

首先我们对X和X_0进行分解：

$$\mathrm{X}=\begin{bmatrix}\mathrm{i}_{N*1} & \mathrm{Z}_{N*(K-1)}\end{bmatrix}$$
$$\mathrm{X}_0=\begin{bmatrix}1 & \mathrm{Z}_{0,1*(K-1)}^{\mathrm{T}}\end{bmatrix}$$

其中

$$\mathrm{i}=\begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}^{\mathrm{T}}$$

代入$(\mathrm{X}^{\mathrm{T}}\mathrm{X})^{-1}$：

$$(\mathrm{X}^{\mathrm{T}}\mathrm{X})^{-1}=\left(\begin{bmatrix}\mathrm{i}^{\mathrm{T}}\\ \mathrm{Z}^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\mathrm{i} & \mathrm{Z}\end{bmatrix}\right)^{-1}=\begin{bmatrix}N & \mathrm{i}^{\mathrm{T}}\mathrm{Z}\\ \mathrm{Z}^{\mathrm{T}}\mathrm{i} & \mathrm{Z}^{\mathrm{T}}\mathrm{Z}\end{bmatrix}^{-1}$$

由分块矩阵的求逆公式可知

$$(\mathrm{X}^{\mathrm{T}}\mathrm{X})^{-1}=\begin{bmatrix}N & \mathrm{i}^{\mathrm{T}}\mathrm{Z}\\ \mathrm{Z}^{\mathrm{T}}\mathrm{i} & \mathrm{Z}^{\mathrm{T}}\mathrm{Z}\end{bmatrix}^{-1}=\begin{bmatrix}\frac{1}{N-\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i}} & -\frac{1}{N}\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}\\ -\frac{1}{N}(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i} & (\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}\end{bmatrix}$$

其中

$$\mathrm{M}_0=\mathrm{I}_{\mathrm{N*N}}-\frac{1}{N}\mathrm{i}\mathrm{i}^{\mathrm{T}}$$
$$\mathrm{M}_0\mathrm{Z}=\begin{bmatrix}\mathrm{Z}_1-\bar{\mathrm{Z}} & \mathrm{Z}_2-\bar{\mathrm{Z}} & \cdots & \mathrm{Z}_{\mathrm{N}}-\bar{\mathrm{Z}}\end{bmatrix}^{\mathrm{T}}$$

将$(\mathrm{X}^{\mathrm{T}}\mathrm{X})^{-1}$代入$var(\hat{y_0})$：

$$var(\hat{y_0})=\sigma^2\begin{bmatrix}1 & {\mathrm{Z}_0}^{\mathrm{T}}\end{bmatrix}(\mathrm{X}^{\mathrm{T}}\mathrm{X})^{-1}\begin{bmatrix}1\\ \mathrm{Z}_0\end{bmatrix}$$
$$=\sigma^2\left(1+\frac{1}{N-\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i}}-\frac{1}{N}\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}\mathrm{Z}_0-\frac{1}{N}{\mathrm{Z}_0}^{\mathrm{T}}(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i}+{\mathrm{Z}_0}^{\mathrm{T}}(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}\mathrm{Z}_0\right)$$

观察上式，括号内每一项都是标量，且$(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}$是定值，于是我们希望将括号内整理成形如$C+\mathrm{P}^{\mathrm{T}}(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}\mathrm{P}$的形式，其中$C$是常数，$\mathrm{P}$与$\mathrm{Z}_0$和$\bar{\mathrm{Z}}$有关。

于是利用配方法可将括号内后三项化为

$$({\mathrm{Z}_0}^{\mathrm{T}}-\frac{1}{N}\mathrm{i}^{\mathrm{T}}\mathrm{Z})(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}(\mathrm{Z}_0-\frac{1}{N}\mathrm{Z}^{\mathrm{T}}\mathrm{i})\equiv({\mathrm{Z}_0}^{\mathrm{T}}-\bar{\mathrm{Z}})(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}({\mathrm{Z}_0}^{\mathrm{T}}-\bar{\mathrm{Z}})^{\mathrm{T}}=$$
$$\underbrace{-\frac{1}{N}\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}\mathrm{Z}_0-\frac{1}{N}{\mathrm{Z}_0}^{\mathrm{T}}(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i}+{\mathrm{Z}_0}^{\mathrm{T}}(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}\mathrm{Z}_0}$$
$$+\frac{\mathrm{i}^{\mathrm{T}}\mathrm{Z}}{N}(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}\frac{\mathrm{Z}^{\mathrm{T}}\mathrm{i}}{N}$$

代入$var(\hat{y_0})$：

$$var(\hat{y_0})=\sigma^2\left(1+\frac{1}{N-\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i}}+(\mathrm{Z_0}^{\mathrm{T}}-\bar{\mathrm{Z}})(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}(\mathrm{Z}_0-\bar{\mathrm{Z}}^{\mathrm{T}})-\underbrace{\frac{\mathrm{i}^{\mathrm{T}}\mathrm{Z}}{N}(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}\frac{\mathrm{Z}^{\mathrm{T}}\mathrm{i}}{N}}\right)$$

由于

$$(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}=(\mathrm{Z}^{\mathrm{T}}\mathrm{Z}-\frac{1}{N}\mathrm{Z}^{\mathrm{T}}\mathrm{i}\mathrm{i}^{\mathrm{T}}\mathrm{Z})^{-1}$$

由两矩阵和的逆的分解公式知

$$\begin{aligned}(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}=(\mathrm{Z}^{\mathrm{T}}\mathrm{Z}-\frac{1}{N}\mathrm{Z}^{\mathrm{T}}\mathrm{i}\mathrm{i}^{\mathrm{T}}\mathrm{Z})^{-1}&=(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}+\frac{(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\frac{1}{N}\mathrm{Z}^{\mathrm{T}}\mathrm{i}\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}}{1-\frac{1}{N}\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i}}\\&=(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}+\frac{(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i}\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}}{N-\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i}}\end{aligned}$$

代入$\frac{\mathrm{i}^{\mathrm{T}}\mathrm{Z}}{N}(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}\frac{\mathrm{Z}^{\mathrm{T}}\mathrm{i}}{N}$：

$$\begin{aligned}\frac{\mathrm{i}^{\mathrm{T}}\mathrm{Z}}{N}(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}\frac{\mathrm{Z}^{\mathrm{T}}\mathrm{i}}{N}&=\frac{\mathrm{i}^{\mathrm{T}}\mathrm{Z}}{N}\left((\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}+\frac{(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i}\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}}{N-\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i}}\right)\frac{\mathrm{Z}^{\mathrm{T}}\mathrm{i}}{N}\\&=\frac{\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i}}{N^2}+\frac{[\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i}][\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i}]}{N(N-\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i})}\end{aligned}$$

由于$\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i}$是标量，可以直接记$\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i}\equiv m$。将上式代入$var(\hat{y_0})$：

$$\begin{aligned}var(\hat{y_0})&=\sigma^2\left(1+\frac{1}{N-\mathrm{i}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{i}}+(\mathrm{Z_0}^{\mathrm{T}}-\bar{\mathrm{Z}})(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}(\mathrm{Z}_0-\bar{\mathrm{Z}}^{\mathrm{T}})-\underbrace{\frac{\mathrm{i}^{\mathrm{T}}\mathrm{Z}}{N}(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}\frac{\mathrm{Z}^{\mathrm{T}}\mathrm{i}}{N}}\right)\\&=\sigma^2\left(1+(\mathrm{Z_0}^{\mathrm{T}}-\bar{\mathrm{Z}})(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}(\mathrm{Z}_0-\bar{\mathrm{Z}}^{\mathrm{T}})+\frac{1}{N-m}-\frac{m}{N^2}+\frac{m^2}{N^2(N-m)}\right)\\&=\sigma^2\left(1+(\mathrm{Z_0}^{\mathrm{T}}-\bar{\mathrm{Z}})(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}(\mathrm{Z}_0-\bar{\mathrm{Z}}^{\mathrm{T}})+\frac{1}{N-m}+\frac{m^2-m(N-m)}{N(N-m)}\right)\\&=\sigma^2\left(1+(\mathrm{Z_0}^{\mathrm{T}}-\bar{\mathrm{Z}})(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}(\mathrm{Z}_0-\bar{\mathrm{Z}}^{\mathrm{T}})+\frac{1}{N-m}-\frac{m}{N(N-m)}\right)\\&=\sigma^2\left(1+(\mathrm{Z_0}^{\mathrm{T}}-\bar{\mathrm{Z}})(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}(\mathrm{Z}_0-\bar{\mathrm{Z}}^{\mathrm{T}})+\frac{N-m}{N(N-m)}\right)\\&=\sigma^2\left(1+\frac{1}{N}+(\mathrm{Z_0}^{\mathrm{T}}-\bar{\mathrm{Z}})(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}(\mathrm{Z}_0-\bar{\mathrm{Z}}^{\mathrm{T}})\right)\end{aligned}$$

其中

$$(\mathrm{Z_0}^{\mathrm{T}}-\bar{\mathrm{Z}})(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}(\mathrm{Z}_0-\bar{\mathrm{Z}}^{\mathrm{T}})=\sum_{i=1}^{K-1}\sum_{j=1}^{K-1}(x_i^0-\bar{x_i})(x_j^0-\bar{x_j})(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})_{ij}^{-1}$$

其中x_i^0代表X_0的第i列元素，$\bar{x_i}$代表X的第i列的均值，二者都是标量。

于是

$$var(\hat{y_0})=\sigma^2\left(1+\frac{1}{N}+\sum_{i=1}^{K-1}\sum_{j=1}^{K-1}(x_i^0-\bar{x_i})(x_j^0-\bar{x_j})(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})_{ij}^{-1}\right)$$

由于$(\mathrm{Z}^{\mathrm{T}}\mathrm{M}_0\mathrm{Z})^{-1}$与$\mathrm{X}_0$无关，显然当$\mathrm{X}_0=\bar{\mathrm{X}}$时，$var(\hat{y_0})$最小，原命题得证。