Stat705 PS 6. Due in class Friday, November 3rd

1. Let $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \mathbf{X}$ is n by p, $p < n$, rank $k \leq p$. Suppose that $\mathbf{X}\boldsymbol{\beta} = \tilde{\mathbf{X}}\mathbf{T}\boldsymbol{\beta} = \tilde{\mathbf{X}}\boldsymbol{\gamma}$ where $\mathbf{T}$ is k by p (rank k), $\boldsymbol{\gamma}$ is length k, and $\tilde{\mathbf{X}}$ is n by k with rank k. We say that $\tilde{\mathbf{X}}\boldsymbol{\gamma}$ is a *reparameterization* of $\mathbf{X}\boldsymbol{\beta}$. Note that $\mathbf{T}\boldsymbol{\beta} = \boldsymbol{\gamma}$ and $\mathbf{X} = \tilde{\mathbf{X}}\mathbf{T}$.

   (a) Prove that $\mathbf{P} = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-}\mathbf{X}^{\mathrm{T}} = \tilde{\mathbf{X}}(\tilde{\mathbf{X}}^{\mathrm{T}}\tilde{\mathbf{X}})^{-}\tilde{\mathbf{X}}^{\mathrm{T}} = \tilde{\mathbf{P}}$. Note that this means $\widehat{\mathbf{y}}$ and $\mathbf{r}$ are the same for $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$ and a reparameterization, $\mathbf{y} = \tilde{\mathbf{X}}\boldsymbol{\gamma} + \mathbf{e}$.

   (b) Suppose $x_i = (c_1, c_1, c_2, c_2, c_3, c_3)^{\mathrm{T}}$ where $c_1 \neq c_2 \neq c_3$. Let $y_i = \beta_0 + \beta_1 1_{x_i=c_2} + \beta_2 1_{x_i=c_3} + e_i, i = 1, \ldots, 6$. Interpret each of the β_ks (in terms of the mean of y_i for particular values of x_i).

      i. Prove that the model above is a reparameterization of

$$\mathbf{y} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} + \mathbf{e}.$$

      ii. Prove that $y_i = \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2 + e_i, i = 1, \ldots, 6$ is a reparameterization too.

      iii. Find a third different reparameterization where one of the parameters can be interpreted as the mean when $x_i = c_2$ minus the mean when $x_i = c_3$.