Problem Set 1

1. Derive the OLS estimator using calculus. To do this first derive $\partial \boldsymbol{a}'\boldsymbol{v}/\partial \boldsymbol{v}$ and $\partial \boldsymbol{v}'\boldsymbol{A}\boldsymbol{v}/\partial \boldsymbol{v}$ for any vector $\boldsymbol{a}$ and any symmetric full rank matrix $\boldsymbol{A}$.

2. Show for the classical multiple regression model that $\hat{\boldsymbol{\beta}}$ and $(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})$ are uncorrelated with $\tilde{\boldsymbol{\beta}}$ being a linear unbiased estimator.

3. Prove that the test statistic
$$\frac{(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)'\boldsymbol{X}'\boldsymbol{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)/k}{s^2} = \frac{(\hat{\boldsymbol{u}}_0'\hat{\boldsymbol{u}}_0 - \hat{\boldsymbol{u}}'\hat{\boldsymbol{u}})/k}{\hat{\boldsymbol{u}}'\hat{\boldsymbol{u}}/(n-k)}$$
for any $\boldsymbol{\beta}_0$ by looking at the geometry of the numerator.

4. Assume $\boldsymbol{u} \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{I})$. Derive the distribution of $\boldsymbol{R}\boldsymbol{\beta}$ with an $m \times k$ matrix with $\text{rank}(\boldsymbol{R}) = m < k$. Construct a scalar test statistic for testing $H_0 : \boldsymbol{R}\boldsymbol{\beta} = \boldsymbol{r}$ and derive its distribution.

5. Consider a model with only two regressors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ and the restriction $\beta_1 + \beta_2 = 4$. Show geometrically the resulting restricted and unrestricted LS estimators.

6. The correct structure of the model is assumed to be $E(\boldsymbol{y}|\boldsymbol{X}) = \boldsymbol{X}\boldsymbol{\beta}$. Since you are not sure about this you run a regression of $\boldsymbol{y}$ on $\boldsymbol{X}$ and $\boldsymbol{W}$. What kind of effects are resulting for the estimates $\hat{\boldsymbol{\beta}}$?

7. Same setting as above. But you only use part of the regressors, i.e. you run a regression of $\boldsymbol{y}$ on $\boldsymbol{X}_1$ with $\boldsymbol{X} = [\boldsymbol{X}_1 \vdots \boldsymbol{X}_2]$. Discuss the effects on $\hat{\boldsymbol{\beta}}_1$.

8. Read the empirical example on individual wages in Verbeek chapter 2. Relate the two problems above to Verbeek's empirical results.

9. Let $\boldsymbol{X}$ and $\boldsymbol{W}$ be two $n \times k$ matrices such that $S(\boldsymbol{X}) \neq S(\boldsymbol{W})$. Show that $\boldsymbol{P} \equiv \boldsymbol{X}(\boldsymbol{W}'\boldsymbol{X})^{-1}\boldsymbol{W}'$ is idempotent but not symmetric. Characterize the spaces that $\boldsymbol{P}$ and $(\boldsymbol{I} - \boldsymbol{P})$ project on to.

10. Show that the estimator $\tilde{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{B}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{B}\boldsymbol{y}$ fulfills the prerequisites of the Gauss-Markov theorem. Derive the covariance matrix of $\tilde{\boldsymbol{\beta}}$.

Problem Set 2

1. Let $Z_i \sim B(1; 0.8)$ and $\bar{Z} = n^{-1} \sum_{i=1}^{n} Z_i$. Compute $E(\bar{Z})$ and $\text{Var}(\bar{Z})$, and the analogue for $\bar{W}$ where $W_i = \exp(Z_i)$ for $n = 1$ and $n = 3$. If you have e.g. EXCEL available perform these calculations also for $n = 10$, $n = 100$, $n = 250$, and $n = 500$, respectively.

    - Relate those results to Jensen's inequality.
    - Explain the asymptotic behavior of the random variable $\exp(\bar{Z})$.
    - How do we see the $\sqrt{n}$ convergence rate?
    - Visualize the probability distribution of $\bar{Z}$.
    - Explain the asymptotic behavior of the random variable $\sqrt{n}\exp(\bar{Z})$.

2. Give answers to Verbeek's Exercise 5.3.

Problem Set 3

1. In a classical linear regression setup a second (artificial) regression is performed as

$$\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{OLS} = \boldsymbol{X}\boldsymbol{b} + \text{Residuals}$$

What is the resulting OLS estimator $\hat{\boldsymbol{b}}$, the sum of squared residuals, and the estimated standard errors of $\hat{\boldsymbol{b}}$?

2. How should such an artificial regression look like for a non-linear regression model

$$\boldsymbol{y} = \boldsymbol{x}(\boldsymbol{\beta}) + \boldsymbol{u}$$

where $\boldsymbol{u}$ has the classical properties?

3. In a linear regression model with autoregressive residuals of order p

$$u_t = \rho_1 u_{t-1} + \ldots + \rho_p u_{t-p} + \varepsilon_t \qquad \varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$$

firstly argue that the model can be interpreted as a non-linear regression model. Secondly, derive an LM-type test for $H_0: \ \rho_1 = \ldots = \rho_p = 0$.

4. In a linear regression model with heteroscedasticity of the form

$$\sigma_t = \sigma^2 h(Z_t \gamma) \qquad \text{with } h(0) = 1$$

derive the score vector and also the information matrix $\boldsymbol{I}(\boldsymbol{\theta})$ with $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\gamma}', \sigma^2)'$. Show that the LM test statistic is equivalent to nR^2 where R^2 is the uncentred R^2 of an auxiliary regression of a vector of ones on the score vectors $\boldsymbol{G}$ where the $t-$th row of $\boldsymbol{G}$ is the individual score contribution $\boldsymbol{g}_t(\boldsymbol{\theta})'$. [Hint: See section 6.2.2. in V].

5. Describe how you would compute the feasible GLS estimator for $\boldsymbol{\beta}$ for the above model with autoregressive residuals. (Do not worry about the initial observations.)