Problem Set no. 1

1. Show that both $\boldsymbol{P_X}$ and $\boldsymbol{M_X}$ are symmetric and idempotent projection matrices.

2. How changing the units of measure of the regressand and regressors in the OLS estimates and their standard errors.

3. Consider the standard linear regression with a single fixed regressor, viz. $y_j = \alpha + \beta x_j + u_j$ with $\boldsymbol{u} \sim \text{IID}N(0, \sigma^2 \mathbf{I}_n)$. Further suppose that all observations are ordered according to increasing values of the regressor and that the sample size n is even. Let then $\bar{x}_1$, $\bar{x}_2$, $\bar{y}_1$ and $\bar{y}_2$ denote the sample mean values of the regressor and regressand over the first and last $n/2$ observations, respectively. Compare the mean squared error of the estimator $\breve{\beta} = \frac{\bar{y}_2 - \bar{y}_1}{\bar{x}_2 - \bar{x}_1}$ to the mean squared error of the OLS estimator.

4. Ramanathan's exercises (pp. 287-291): 10.1($a \rightarrow g$), 10.2→10.8, 10.10

Problem Set no. 2

1. Show that the t-statistic is simply $(n-k)^{1/2}\cot\phi$ and then interpret the cases in which $\phi = 0°$ and $\phi = 90°$. Now derive the order of magnitude of the t-statistic for $0° < \phi < 90°$ to demonstrate that the t-statistic entails a consistent test, that is to say, is such that the probability of rejecting the null hypothesis converges to one as the sample size grows.

2. Consider the following data generating mechanism $\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{u}$, where $\boldsymbol{u} \sim N(\mathbf{0}, \sigma_0^2\mathbf{I}_n)$. Consider the following quantities: $\hat{\boldsymbol{\beta}} = \begin{bmatrix}\hat{\boldsymbol{\beta}}_1^\top & \hat{\boldsymbol{\beta}}_2^\top\end{bmatrix}^\top$ and $\tilde{\boldsymbol{\beta}} = \begin{bmatrix}\tilde{\boldsymbol{\beta}}_1^\top & \mathbf{0}^\top\end{bmatrix}^\top$. The first corresponds to the OLS estimator of the unrestricted model, whereas the second is the OLS estimator of the restricted regression model $\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{u}$. Which estimator performs better in terms of the mean squared error matrix?

3. Consider now the converse problem in which the data generating mechanism reads $\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{u}$, where $\boldsymbol{u} \sim N(\mathbf{0}, \sigma_0^2\mathbf{I}_n)$. What does happen with $\hat{\boldsymbol{\beta}} = \begin{bmatrix}\hat{\boldsymbol{\beta}}_1^\top & \hat{\boldsymbol{\beta}}_2^\top\end{bmatrix}^\top$ and $\tilde{\boldsymbol{\beta}} = \begin{bmatrix}\tilde{\boldsymbol{\beta}}_1^\top & \mathbf{0}^\top\end{bmatrix}^\top$ in terms of consistency and relative efficiency?

4. Illustrate the fact that unbiasedness neither implies nor is implied by consistency by investigating the properties of the following estimators of the mean μ of a population with variance σ^2: $\tilde{\mu} = \frac{1}{n-k}\sum_{j=1}^n x_j$, $\breve{\mu} = \frac{1}{k}\,x_1 + \frac{k-1}{k(n-1)}\sum_{j=2}^n x_j$, and $\bar{\mu} = \frac{k}{n}\sum_{j=1}^n x_j$ for some constant $k \neq 1$.

Problem Set no. 3

1. Consider the regression model $\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{u}$ with fixed regressors and $E(\boldsymbol{u}|\boldsymbol{X}_1, \boldsymbol{X}_2) = \boldsymbol{X}_1\boldsymbol{\gamma}$. Compute the bias of the OLS estimator $\hat{\boldsymbol{\beta}}$.

2. Consider the regression model $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}$ with fixed regressors and $E(\boldsymbol{u}\boldsymbol{u}^\top) = \Omega$, where Ω is a symmetric and positive definite matrix. Compute the LM and Wald tests for $r(\boldsymbol{\beta}) = 0$.

3. Consider the simple linear regression $\boldsymbol{y} = \beta\boldsymbol{x} + \boldsymbol{u}$, where both $\boldsymbol{y}$ and $\boldsymbol{x}$ are demeaned, and $\boldsymbol{u} \sim \text{IID}N(0, \sigma^2\mathbf{I}_n)$. The fitted values are then $\hat{\boldsymbol{y}} \equiv \hat{\beta}\boldsymbol{x}$. Imagine that one wishes to estimate the value x_0 that could have given rise to y_0 and obtain a confidence interval for x_0.

4. Consider two nonnested linear regressions $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}$ and $\boldsymbol{y} = \boldsymbol{Z}\boldsymbol{\gamma} + \boldsymbol{v}$, where $E(\boldsymbol{u}\boldsymbol{u}^\top) = \sigma_u^2\mathbf{I}_n$ and $E(\boldsymbol{v}\boldsymbol{v}^\top) = \sigma_v^2\mathbf{I}_n$. The J-test by Davidson and MacKinnon hinges on testing whether $\alpha = 0$ in the the artificial nesting given by $\boldsymbol{y} = (1-\alpha)\boldsymbol{X}\boldsymbol{\beta} + \alpha\boldsymbol{Z}\boldsymbol{\gamma} + \boldsymbol{e}$. To avoid the lack of joint identifiability of $(\alpha, \boldsymbol{\beta}, \boldsymbol{\gamma})$, it suffices to replace $\boldsymbol{Z}\boldsymbol{\gamma}$ with $\boldsymbol{Z}\hat{\boldsymbol{\gamma}} \equiv \boldsymbol{P}_{\boldsymbol{Z}}\boldsymbol{y}$. Demonstrate that the P-test based on the t-statistic for $a = 0$ from the Gauss-Newton regression $\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{y} = \boldsymbol{X}\boldsymbol{b} + a(\boldsymbol{Z}\hat{\boldsymbol{\gamma}} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) + \text{residuals}$ yields precisely the same result. Define the implicit null and alternative hypotheses of the J- and P-tests.

5. Consider the standard linear regressions $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}$. Use the Gauss-Newton approach to derive a test whether $E|u_j| = h(\alpha + \boldsymbol{Z}_j\boldsymbol{\gamma})$.

Problem Set no. 4

1. Consider the nonlinear regression $\boldsymbol{y} = \boldsymbol{x}(\boldsymbol{\beta}) + \boldsymbol{u}$ with $\boldsymbol{u} \sim \mathrm{IID}(o, \sigma^2 \mathbf{I}_n)$. The corresponding sum-of-squares function is

   $$SSR(\boldsymbol{\beta}) = \big(\boldsymbol{y} - \boldsymbol{x}(\boldsymbol{\beta})\big)^\top \big(\boldsymbol{y} - \boldsymbol{x}(\boldsymbol{\beta})\big).$$

   Differentiate this expression with respect to the components of the k-vector $\boldsymbol{\beta}$ and set all partial derivatives to zero so as to derive the first-order conditions (i.e. normal equations) that the nonlinear least squares (NLS) estimator $\hat{\boldsymbol{\beta}}$ must satisfy.

2. Suppose the asymptotic identification of the nonlinear regression model by the sum-of-squares function, that is to say, $SSR(\boldsymbol{\beta}_0) \neq SSR(\boldsymbol{\beta})$. Show that the above NLS estimator is consistent provided that the sequence $\left\{n^{-1} \sum_{j=1}^{n} x_j(\boldsymbol{\beta}) u_j\right\}$ satisfies a weak uniform law of large numbers with probability limit of zero for all $\boldsymbol{\beta} \in \Theta$ and the probability limit of the sequence $\left\{n^{-1} \sum_{j=1}^{n} x_j(\boldsymbol{\beta}) x_j(\boldsymbol{\beta}')\right\}$, for any $\boldsymbol{\beta}' \in \Theta$, is finite, continuous in $\boldsymbol{\beta}$ and $\boldsymbol{\beta}'$, nonstochastic, and uniform with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\beta}'$.

3. Consider an asymptotically identified nonlinear regression that results in a consistent NLS estimator. Let $\boldsymbol{H}_j(y_j, \boldsymbol{\beta}) \equiv \mathrm{D}_{\beta\beta}\big(y_j - x_j(\boldsymbol{\beta})\big)^2$ denote the Hessian of the contribution to the sum-of-squares function from the observation j, and assume that the sequence $\left\{\frac{1}{n} \sum_{j=1}^{n} \boldsymbol{H}_j(y_j, \boldsymbol{\beta})\right\}$ satisfies a weak uniform law of large numbers for $\boldsymbol{\beta}$ in the vicinity of $\boldsymbol{\beta}_0$. Suppose further that the sequence $\left\{n^{-1} \sum_{j=1}^{n} \boldsymbol{X}_j^\top(\boldsymbol{\beta}) u_j\right\}$, where the $n \times k$ matrix $\boldsymbol{X}^\top(\boldsymbol{\beta})$ has typical element $\boldsymbol{X}_{jk}(\boldsymbol{\beta}) \equiv \frac{\partial}{\partial \beta_k} x_j(\boldsymbol{\beta})$, satisfies a central limit theorem. Show the asymptotic normality of $n^{1/2}\big(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\big)$ with variance $\sigma_0^2 \ \mathrm{plim}_{n\to\infty}(n^{-1}\boldsymbol{X}_0^\top \boldsymbol{X}_0)^{-1}$, where $\boldsymbol{X}_0 \equiv \boldsymbol{X}(\boldsymbol{\beta}_0)$.

   Hint: The consistency of the NLS estimator ensures that $\hat{\boldsymbol{\beta}}$ is close to $\boldsymbol{\beta}_0$, hence one may expand the normal equations in a short Taylor expansion around $\boldsymbol{\beta}_0$.

Problem Set no. 5

1. Show that the projection matrices $\boldsymbol{P}_{\boldsymbol{X}}^{\Omega} \equiv \boldsymbol{X}\left(\boldsymbol{X}^{\top}\Omega^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\top}\Omega^{-1}$ and $\boldsymbol{M}_{\boldsymbol{X}}^{\Omega} \equiv \mathbf{I}_n - \boldsymbol{P}_{\boldsymbol{X}}^{\Omega}$ are idempotent, though not necessarily symmetric.

2. Demonstrate that the Kruskal's theorem indeed holds.

3. Consider the linear regression model $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}$, where $E\left(\boldsymbol{u}\boldsymbol{u}^{\top}\middle|\boldsymbol{X}\right) = \sigma^2\Delta$, with σ^2 unknown but Δ is a $n \times n$ symmetric and positive definite known matrix. Under normality, the log-likelihood function is

$$\ell(\boldsymbol{y};\boldsymbol{\theta}) = -\frac{n}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2}\log|\Delta| - \frac{1}{2\sigma^2}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^{\top}\Delta^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}),$$

   where $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2, \Delta)$. Compute the concentrated log-likelihood function with respect to σ^2.

4. Give an intuitive explanation of why the feasible GLS is asymptotic equivalent to the genuine GLS using the fact that $\tilde{\alpha}$ is a root-n consistent estimator of α and that $\Omega(\alpha)$ is differentiable at α.

5. Consider the linear regression model $y_j = \boldsymbol{X}_j\boldsymbol{\beta} + u_j$ with $E\left(\boldsymbol{u}\boldsymbol{u}^{\top}\middle|\boldsymbol{X}\right) = \Omega$, where Ω is $n \times n$ diagonal matrix with elements ω_j^2. Suppose that we ignore the heteroskedastic nature of the regression function and estimate the parameter vector $\boldsymbol{\beta}$ by OLS. The resulting covariance matrix is simply $\sigma^2(\boldsymbol{X}^{\top}\boldsymbol{X})^{-1}$, though we know that the correct covariance matrix of the OLS estimator is $(\boldsymbol{X}^{\top}\boldsymbol{X})^{-1}\boldsymbol{X}^{\top}\Omega\boldsymbol{X}(\boldsymbol{X}^{\top}\boldsymbol{X})^{-1}$. Show that σ^2 is the probability limit of the average of the ω_j^2's. What happens in the event that $\boldsymbol{X}$ consists solely of a constant term?

Problem Set no. 6

1. Consider the linear regression model $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}$, where u_t follows an ARMA(1,1) process, i.e., $u_t = \rho u_{t-1} + \epsilon_t + \theta\epsilon_{t-1}$ and $\epsilon_t \sim \text{IID}(0, \omega^2)$. Compute the covariance matrix of $\boldsymbol{u}$.

2. Consider the linear regression model $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}$, where u_t follows an AR(1) process, that is to say, $u_t = \rho u_{t-1} + \epsilon_t$ and $\epsilon_t \sim \text{IID}(0, \omega^2)$. Show how to test this particular model against the alternative model $y_t = \boldsymbol{X}_t\boldsymbol{\beta} + \rho y_{t-1} + \boldsymbol{X}_{t-1}\gamma + \epsilon_t$, where $\epsilon_t \sim \text{IID}(0, \omega^2)$, using an F-test and then compare with a test based on the appropriate Gauss-Newton regression.

3. Consider the linear regression model $\boldsymbol{y} = \beta_1\boldsymbol{x}_1 + \beta_2\boldsymbol{x}_2 + \boldsymbol{u}$, where $\boldsymbol{u} \sim \text{IID}(0, \sigma^2)$ and both β_1 and β_2 are scalars. Assuming that only $\boldsymbol{x}_1$ is exogenous, derive the asymptotic properties of the OLS estimator of $\theta \equiv \left(\beta_1, \beta_2, \sigma^2\right)$.

4. Let $\breve{y}_t = y_t/\breve{u}_t$, $\breve{\boldsymbol{X}}_t = \boldsymbol{X}_t/\breve{u}_t$ and $\boldsymbol{Z}_t = \boldsymbol{X}_t\breve{u}_t$, where $\breve{u}_t^2$ denotes some estimate of u_t^2: either $\hat{u}_t^2$, $\frac{n}{n-k}\,\hat{u}_t^2$, $\frac{1}{1-h_t}\,\hat{u}_t^2$, or $\frac{1}{(1-h_t)^2}\,\hat{u}_t^2$. Show that regressing $\breve{y}_t$ on $\breve{\boldsymbol{X}}_t$ using $\boldsymbol{Z}_t$ as the vector of instruments yields the same estimates as an OLS regression of y_t on $\boldsymbol{X}_t$. Demonstrate further that the IV covariance matrix is proportional to the HCCME corresponding to the set of residuals $\breve{u}_t$. Lastly, show that the constant of proportionality is just s^2, the IV estimator of the error variance, which should converge to unity as the sample size grows.

5. Show that the difference between the sum of squared residuals of the second-stage regression of the 2SLS estimator for the models $\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{u}$ and $\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{u}$, where $\boldsymbol{u} \sim \text{IID}(\boldsymbol{0}, \sigma_0^2\mathbf{I}_n)$, is

$$\boldsymbol{u}'\boldsymbol{M}_{\boldsymbol{P_W X}_1}\boldsymbol{P_W}\boldsymbol{X}_2\left(\boldsymbol{X}_2'\boldsymbol{P_W}\boldsymbol{M}_{\boldsymbol{P_W X}_1}\boldsymbol{P_W}\boldsymbol{X}_2\right)^{-1}\boldsymbol{X}_2'\boldsymbol{P_W}\boldsymbol{M}_{\boldsymbol{P_W X}_1}\boldsymbol{u}.$$

Further, show that $n^{-1/2}\boldsymbol{X}_2\boldsymbol{P_W}\boldsymbol{M}_{\boldsymbol{P_W X}_1}\boldsymbol{u}$ is asymptotically normally distributed with covariance matrix

$$\sigma_0^2 \operatorname*{plim}_{n\to\infty}\left(\frac{\boldsymbol{X}_2'\boldsymbol{P_W}\boldsymbol{M}_{\boldsymbol{P_W X}_1}\boldsymbol{P_W}\boldsymbol{X}_2}{n}\right)^{-1}.$$