$$SSR = \sum_{t=1} u_t^2 = u'u = (Y - X\beta)^T(Y - X\beta) = Y^TY - Y^TX\beta - \beta^TX^TY$$

$$= Y'Y - 2\beta^TX^TY + \beta^TX^TX\beta$$

Therefore, we set the first derivatives to zero:

$$\partial SSR/\partial\beta = \partial/\partial\beta\,(Y^TY - 2\beta^TX^TY + \beta^TX^TX\beta) = -2X^TY + 2X^TX\beta =$$

Solving for β, we obtain:

$$\hat{\beta} = (X^TX)^{-1}X^TY$$

b) What is the matrix that, if post-multiplied by Y, returns the fitted value $\hat{Y} =$
the geometric interpretation of such operation? (3 points)
It is the orthogonal projection matrix $P_X = X(X^TX)^{-1}X^T$, which is the projects any vector onto the space spanned by the columns of X. The proj that the residual $\hat{u} = Y - \hat{Y}$ is orthogonal to $\hat{Y}$.

c) What happens to the fitted value and to the OLS estimates if, instead of co matrix of regressors X, I consider a linear transformation $\tilde{X} = XD$, w nonsingular matrix? (3 points)
The fitted value remains the same (and therefore also the residuals rema while the new OLS estimate become $\tilde{\beta} = D^{-1}\hat{\beta}$.

d) Derive the variance of the OLS estimator if we know that $Var(u) = \Sigma$, wh matrix, and that $E(u|X) = E(u) = 0$. (3 points)
We know that the OLS estimator can be written as: $\hat{\beta} = \beta + (X^TX)^{-1}X^Tu$ β (given that $E(u|X) = E(u) = 0$). Therefore, its variance is

$$Var(\hat{\beta}) = E\left[(\hat{\beta} - \beta)(\hat{\beta} - \beta)^T\right] = E\left\{(X^TX)^{-1}X^Tu\left[(X^TX)^{-1}X^T\right.\right.$$

$$= E\left[(X^TX)^{-1}X^Tuu^TX(X^TX)^{-1}\right] = (X^TX)^{-1}X^TE(uu^T)X(X^TX$$

$$= (X^TX)^{-1}X^T\Sigma X(X^TX)^{-1}$$

e) Under what assumptions is the OLS estimator BLUE? What does BLUE mear
The OLS estimator is the Best Linear Unbiased Estimator, i.e. the lin estimator with the smallest variance, if the following assumptions hold:
1) **The DGP is $Y = X\beta + u$;**
2) **X has full column rank;**
3) **The regressors are exogenous: $E(u|X) = E(u) = 0$;**
4) **Errors are not serially correlated: Σ is diagonal;**
5) **Errors are homoscedastic: the diagonal elements of Σ are all equa**

2. Imagine that data on working men was used to estimate the parameters of the follov (these results are made up):

$$educ = \beta_0 + \beta_1 * sibs + \beta_2 * meduc + \beta_3 * feduc + \beta_4 * city + u$$