

# Econometrics II

Fabian Waldinger (LMU Munich)

# Lecture Structure

- ① Review of OLS with Matrix algebra
- ② Example of OLS with one regressor
- ③ Second Order Condition and what it implies
- ④ Gauss-Markov Assumptions
- ⑤ Unbiasedness of OLS
- ⑥ Variance-Covariance Matrix of the OLS estimator

# Recap from Last Lecture

- Linear Model:

$$y = \mathbf{X}\beta + \varepsilon$$

- We find the  $\hat{\beta}$  that minimizes the RSS
- FOC:

$$-2\mathbf{X}'\hat{\varepsilon} = 0$$

$$\mathbf{X}'\hat{\varepsilon} = 0 \rightarrow \mathbf{X}'[y - \mathbf{X}\hat{\beta}] = 0 \rightarrow \hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y$$

- For this last step we have to be able to invert  $\mathbf{X}'\mathbf{X}$
- From the SOC we know that  $\mathbf{X}'\mathbf{X}$  is positive definite  $\rightarrow$  it is invertible
- $\mathbf{X}'\mathbf{X}$  has to have full rank to be invertible (i.e. it has rank  $k$ )

# Rank of $\mathbf{X}'\mathbf{X}$

- $\mathbf{X}'\mathbf{X}$  has rank  $k$  if  $\mathbf{X}$  has rank  $k$
- For  $\mathbf{X}$  to have rank  $k$  we require that there is no perfect linear combination between the columns of  $\mathbf{X}$ ; e.g. we do not want the following  $\mathbf{X}$  matrix (or other similar matrices)
  
- This condition ensures that there is no perfect multicollinearity
- This requires that  $N$  is at least as large as  $k$  (because the rank is at most the minimum of the two dimensions)

## Example: Only One Regressor

- We now study the example of only one regressor ( $y_i = \beta_1 + \beta_2 x_{i2} + \varepsilon_i$ ). In that case:

- Hence  $\mathbf{X}'\mathbf{X} =$

## Example: Only one Regressor

- Hence the OLS estimator  $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y$  is:

## Example: Only one Regressor

- What is  $(\mathbf{X}'\mathbf{X})^{-1}$ ?
- For a  $2 \times 2$  matrix:
  
  
  
  
  
  
  
  
  
  
- The inverse is defined as:

## Example: Only one Regressor

- Hence  $(\mathbf{X}'\mathbf{X})^{-1}$  is:

$$\frac{1}{N \sum_{i=1}^N x_{i2}^2 - \sum_{i=1}^N x_{i2} \sum_{i=1}^N x_{i2}} \begin{pmatrix} \sum_{i=1}^N x_{i2}^2 & - \sum_{i=1}^N x_{i2} \\ - \sum_{i=1}^N x_{i2} & N \end{pmatrix}$$

- This can be simplified to:

$$\frac{1}{N \sum_{i=1}^N x_{i2}^2 - N^2 \bar{x}_2^2} \begin{pmatrix} \sum_{i=1}^N x_{i2}^2 & -N\bar{x}_2 \\ -N\bar{x}_2 & N \end{pmatrix}$$



## Example: Only one Regressor

- Hence  $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y$  is:

$$\frac{1}{N \sum_{i=1}^N x_{i2}^2 - N^2 \bar{x}_2^2} \begin{pmatrix} \sum_{i=1}^N x_{i2}^2 & -N\bar{x}_2 \\ -N\bar{x}_2 & N \end{pmatrix} \begin{pmatrix} N\bar{y} \\ \sum_{i=1}^N x_{i2}y_i \end{pmatrix}$$

$$= \frac{1}{N \sum_{i=1}^N x_{i2}^2 - N^2 \bar{x}_2^2} \begin{pmatrix} \sum_{i=1}^N x_{i2}^2 N\bar{y} - N\bar{x}_2 \sum_{i=1}^N x_{i2}y_i \\ -N^2 \bar{x}_2 \bar{y} + N \sum_{i=1}^N x_{i2}y_i \end{pmatrix}$$

## Example: $\hat{\beta}_2$

- Hence

$$\hat{\beta}_2 = \frac{-N^2 \bar{x}_2 \bar{y} + N \sum_{i=1}^N x_{i2} y_i}{N \sum_{i=1}^N x_{i2}^2 - N^2 \bar{x}_2^2} = \frac{\frac{1}{N} \sum_{i=1}^N x_{i2} y_i - \bar{x}_2 \bar{y}}{\frac{1}{N} \sum_{i=1}^N x_{i2}^2 - \bar{x}_2^2}$$

# The Second Order Condition Simple Regression

- We will look at the SOC in the simple linear regression case (only one  $x$  variable)
- Recall the SOC from the last lecture:  $2(\mathbf{X}'\mathbf{X})$
- We therefore have to understand whether  $(\mathbf{X}'\mathbf{X})$  is positive definite (would imply a minimum of the RSS) or negative definite (would imply a maximum)
- In the simple linear regression example:

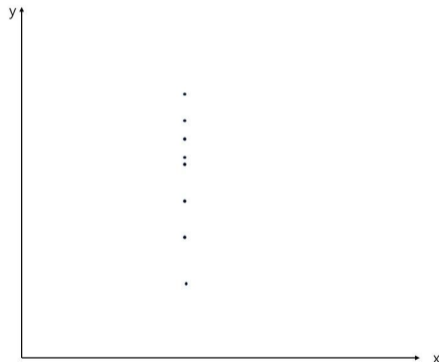
$$(\mathbf{X}'\mathbf{X}) = \begin{pmatrix} N & \sum_{i=1}^N x_{i2} \\ \sum_{i=1}^N x_{i2} & \sum_{i=1}^N x_{i2}^2 \end{pmatrix}_{2 \times 2}$$

# The Second Order Condition Simple Regression

- How to check whether  $(\mathbf{X}'\mathbf{X})$  is positive or negative definite?
- Recall that a 2x2 matrix  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  is positive definite iff the two leading principle minors are positive, i.e.:
  - ①  $a > 0$
  - ②  $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} > 0$
- In our example:
  - ①  $a = N > 0$
  - ②  $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{N} \sum_{i=1}^N x_{i2}^2 - \bar{x}^2 > 0$
- The latter expression is the expression of the sample variance of  $x$  which has to be  $> 0$

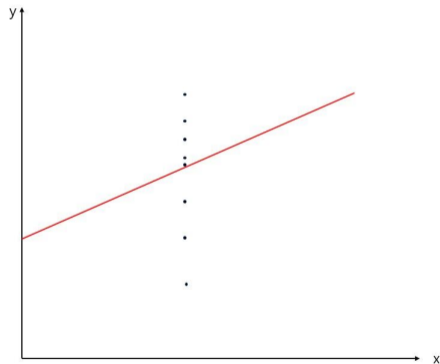
# What Would Happen if $s_{x^2} = 0$ ?

- What would happen if  $s_{x^2} = 0$ ?
- The data would look as follows:



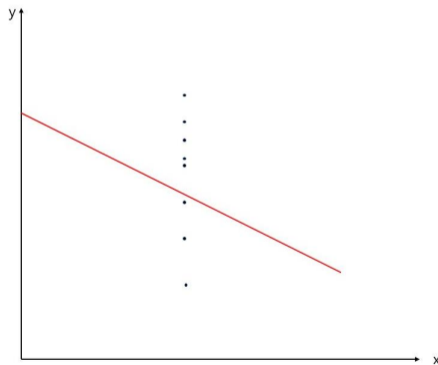
# What Would Happen if $s_{x^2} = 0$ ?

- What would happen if  $s_{x^2} = 0$ ?
- The data would look as follows:



# What Would Happen if $s_{x^2} = 0$ ?

- What would happen if  $s_{x^2} = 0$ ?
- The data would look as follows:



# What Would Happen if $s_{x^2} = 0$ ?

- If there was no variation in  $X$  we could not estimate a regression line
- This would be a violation of SLR.3 from last term
- Here we see mathematically why SLR.3 needs to hold



# Gauss-Markov Assumptions

- In the lectures with Joachim Winter you covered the Gauss-Markov assumptions
- We now look at GM assumptions that have been adapted for the use of matrix Algebra
  - ① The true model is linear in parameters:

$$y = \mathbf{X}\beta + \varepsilon$$

- ① No Perfect Collinearity  
The matrix  $X$  has rank  $k$
- ② Zero Conditional Mean

$$E(\varepsilon|\mathbf{X}) = 0$$

- Note: assumption 3 is implied by MLR.4. under the random sampling assumption MLR.2 from last term

# Unbiasedness of OLS Estimator

- These 3 assumptions need to hold to ensure that the OLS estimator is unbiased
- A particular focus is on GM assumption 3:  $E(\varepsilon|X) = 0$
- This assumption ensures that the value of  $X$  does not tell us anything about the mean of  $\varepsilon$  (but may tell us something about the variance...)
- There are also stronger or weaker versions of GM assumption 3
- Under stronger assumptions OLS is still unbiased and under the weaker of assumption [ $\text{corr}(\varepsilon_i, x_{ik}) = 0$ ] we can show that OLS is consistent (but not unbiased)

# Proof that OLS is Unbiased

- We now show that the OLS estimator is unbiased if GM1-GM3 hold:

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y$$

- We now plug in the true model for  $y$ :

- This can be rewritten as:

# Proof that OLS is Unbiased

- What is the expected value of the expression on the previous slide?
- We use assumption GM3

- The last step follows from GM3
- This shows that  $\hat{\beta}_{OLS}$  is an unbiased estimator for  $\beta$

# Variance-Covariance Matrix of the Error Term

- The variance-covariance matrix of  $\varepsilon|\mathbf{X}$  is defined as follows:

# GM Assumption 4

- We now introduce GM assumption 4:

$$\text{Var}(\varepsilon|\mathbf{X}) = \sigma^2 I$$

- This is short for:

- This assumptions implies two important characteristics:
  - ① No Heteroscedasticity
  - ② No Autocorrelation

# The Variance-Covariance Matrix of the OLS Estimator

- Under GM1-GM4 we can derive the variance-covariance matrix of the OLS estimator
- Under GM1-GM2 we have shown that:

$$\hat{\beta} = \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\varepsilon$$

$$\rightarrow (\hat{\beta} - \beta) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\varepsilon$$

# The Variance-Covariance Matrix of the OLS Estimator

- What is the variance-covariance matrix of  $\hat{\beta}$  ?
- The above is just the formula for the variance-covariance matrix in matrix notation
- From assumptions GM1-GM3 we know that  $E[\hat{\beta}|\mathbf{X}] = \beta$ , hence:
- Now we substitute  $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\varepsilon$  for  $\hat{\beta} - \beta$



# The Variance-Covariance Matrix of the OLS Estimator

- Now using matrix algebra:
  
  
  
  
  
  
  
  
  
  
- Because  $(\mathbf{X}'\mathbf{X})^{-1}$  is symmetric:  $(\mathbf{X}'\mathbf{X})^{-1'} = (\mathbf{X}'\mathbf{X})^{-1}$

# The Variance-Covariance Matrix of the OLS Estimator

- Because  $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$  and  $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$  are fixed with respect to  $\mathbf{X}$  we can rewrite the expression on the previous slide as:

$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E\{\varepsilon\varepsilon'|\mathbf{X}\}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

- Because of GM4  $E\{\varepsilon\varepsilon'|\mathbf{X}\} = \sigma^2 I$ , hence the expression above simplifies to:

$$\begin{aligned} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\sigma^2 I\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\ &= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'I\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\ &= \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \end{aligned}$$

- The last expression is the formula for the variance-covariance matrix of the OLS estimator