QBUS6830: Financial Time Series and Forecasting
Matrix Algebra and Regression

Semester 1, 2017

Introduction

Tsay, Chapter 8, Appendix A

Brooks, Appendix 1 Other resources:

Magnus and Neudecker (2007) ”Matrix Differential Calculus with Applications in Statistics and Econometrics”

There are many free online resources available., e.g.

http://matrixcookbook.com/

http://www.ssc.wisc.edu/ bhansen/econometrics/ (see Appendix A)

An $m \times n$ matrix $\boldsymbol{A}$ is a rectangular array of numbers,

$$\boldsymbol{A} := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Alternative notation $\boldsymbol{A} = \{a_{ij}\}$.

Background and motivation

Systems of linear equations (simultaneous equations)

$$\begin{aligned} 2x_1 + 3x_2 &= 1 \\ 3x_1 - 2x_2 &= 2 \end{aligned}$$

can also be represented as a matrix and vector system:

$$\begin{pmatrix} 2 & 3 \\ 3 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

or in general:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

the general (exact) solution to this system is:

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$$

when $\boldsymbol{A}^{-1}$ exists. Under what conditions does it exist??

Operations

- addition: If $\boldsymbol{A}$ is an $m \times n$ matrix and $\boldsymbol{B}$ a $m \times n$ matrix, then

$$\boldsymbol{A} + \boldsymbol{B} = \{a_{ij} + b_{ij}\}.$$

  Note: this can ONLY be performed if $\boldsymbol{A}$ and $\boldsymbol{B}$ have EXACTLY the same dimensions
- i.e.

$$\begin{pmatrix} 2 & 3 \\ 3 & -2 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 2+1 & 3+2 \\ 3+3 & -2+4 \end{pmatrix} = \begin{pmatrix} 3 & 5 \\ 6 & 2 \end{pmatrix}$$

- But ...

$$\begin{pmatrix} 2 & 3 \\ 3 & -2 \end{pmatrix} + \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \text{undefined}$$

Operations

- multiplication: If $\boldsymbol{A}$ is an $m \times n$ matrix and $\boldsymbol{B}$ a $n \times q$ matrix, then

$$\boldsymbol{AB} = \left\{ \sum_{j=1}^{n} a_{ij} b_{jk} \right\}.$$

Note, often $\boldsymbol{AB} \neq \boldsymbol{BA}$ (only sometimes could these be equal, see below) .

- Thus:

$$\begin{pmatrix} 2 & 3 \\ 3 & -2 \end{pmatrix} \times \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 2\times 1 + 3 \times 3 & 2 \times 2 + 3 \times 4 \\ 3 \times 1 - 2 \times 3 & 3 \times 2 - 2 \times 4 \end{pmatrix}$$

$$= \begin{pmatrix} 11 & 16 \\ -3 & -2 \end{pmatrix}$$

- Each row of $\boldsymbol{A}$ is ”poured” down each column of $\boldsymbol{B}$

Operations

- multiplication: Also

$$\begin{pmatrix} 2 & 3 \\ 3 & -2 \end{pmatrix} \times \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \\ -3 \end{pmatrix}$$

- But ...

$$\begin{pmatrix} 1 \\ 3 \end{pmatrix} \times \begin{pmatrix} 2 & 3 \\ 3 & -2 \end{pmatrix} = \text{undefined}$$

- multiplication ($\boldsymbol{AB}$) is ONLY defined when the column dimension of $\boldsymbol{A}$ EQUALS the row dimension of $\boldsymbol{B}$.

Operations

- transpose:

$$\boldsymbol{A}' = (\widetilde{a}_{ij}),\ \widetilde{a}_{ij} = a_{ji}.$$

- thus

$$\begin{pmatrix} 2 & 1 \\ 3 & -2 \end{pmatrix}' = \begin{pmatrix} 2 & 3 \\ 1 & -2 \end{pmatrix}$$

- Useful: $(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}'$
- Note that the reverse order is required since if $\boldsymbol{A}$ has the same column dimension as $\boldsymbol{B}$'s row dimension then $\boldsymbol{B}'$ has the same COLUMN dimension as the ROW dimension of $\boldsymbol{A}'$
- i.e. if $\boldsymbol{AB}$ exists, then $\boldsymbol{B}'\boldsymbol{A}'$ also exists, but $\boldsymbol{A}'\boldsymbol{B}'$ may not!

Vectors

$$\boldsymbol{x} := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$$

Inner product: $\boldsymbol{x}'\boldsymbol{x} = \sum x_i^2$

Outer product: $\boldsymbol{x}\boldsymbol{x}'$ is an $m \times m$ matrix(!) which is:

Regression models

Recall the SLR model

$$y_1 = \beta_0 + \beta_1 x_1 + u_1$$

$$\vdots$$

$$y_n = \beta_0 + \beta_1 x_n + u_n$$

It can be rewritten in a compact way

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u},$$

where

$$\boldsymbol{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \boldsymbol{X} = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}, \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}, \boldsymbol{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}.$$

Much of the LS and sampling distribution theory becomes EASIER this way !

Regression ctd

Note that

$$\boldsymbol{X}\boldsymbol{\beta} = \begin{pmatrix} \beta_0 + \beta_1 x_1 \\ \beta_0 + \beta_1 x_2 \\ \vdots \\ \beta_0 + \beta_1 x_n \end{pmatrix}$$

as required

Identity

- identity matrix

$$\boldsymbol{I}_n := \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}.$$

  This is a matrix version of 1
- i.e. when these multiplications exist: $\boldsymbol{A}\boldsymbol{I} = \boldsymbol{A} = \boldsymbol{I}\boldsymbol{A}$
- this is a square matrix (i.e. column dimension = row dimension)
- So, e.g.

$$\begin{pmatrix} 2 & 1 \\ 3 & -2 \end{pmatrix} \times \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} =$$

  Try it!

Inversion

- inverse: A matrix $\boldsymbol{A}$ has an inverse, $\boldsymbol{A}^{-1}$, if and only if

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = I$$

- Only square matrices can have an inverse
- Why can't

$$\boldsymbol{A} = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \end{pmatrix}$$

have an inverse??

- What about $\boldsymbol{A}'$??

Regression estimation

- Imagine the SLR model

  $$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u},$$

  with 0 error, $\boldsymbol{u} = \boldsymbol{0}$.
- $$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}$$

  implies that

  $$\boldsymbol{\beta} = \boldsymbol{X}^{-1}\boldsymbol{y}$$

- Is this correct? Is it possible?
- Recall: only square matrices can have an inverse.
- What is the dimension of $\boldsymbol{X}$? Why might it preclude $\boldsymbol{X}^{-1}$ existing?

Square matrices: symmetry and inversion rules

- symmetric matrices are such that: $\boldsymbol{A}' = \boldsymbol{A}$
- Only square matrices can be symmetric!
- Matrix inversion has nice properties: $(\boldsymbol{A}^{-1})' = (\boldsymbol{A}')^{-1}$ and $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$
- This last result must be true since $\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}\boldsymbol{AB} = I$ and $\boldsymbol{ABB}^{-1}\boldsymbol{A}^{-1} = I$.
- Note that $\boldsymbol{I}$ is always symmetric and square.

Square matrices: inversion rules

- determinant: A square matrix $\boldsymbol{A}$ has a determinant, denoted $|\boldsymbol{A}|$.

- A matrix $\boldsymbol{A}$ has a determinant only if it is square.
- A square matrix $\boldsymbol{A}$ has an inverse if and only if $|\boldsymbol{A}| \neq 0$.
- $|\boldsymbol{A}|$ can only equal 0 if
  1. any column (or row) of $\boldsymbol{A}$ is a linear combination of two or more columns (or rows) of $\boldsymbol{A}$.
  2. a row or column of $\boldsymbol{A}$ is entirely full of 0s.

OLS

OLS estimator minimizes the sum of squared residuals, which can be written as

$$\sum u_i^2 = \boldsymbol{u}'\boldsymbol{u} = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})$$

or

$$\boldsymbol{y}'\boldsymbol{y} - \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}$$

To minimise this, we need to know how to differentiate vectors and matrices ...

First, note that, for two equal dimension vectors $\boldsymbol{a}, \boldsymbol{b}$:

$$(\boldsymbol{a} + \boldsymbol{b})'(\boldsymbol{a} + \boldsymbol{b}) = \boldsymbol{a}'\boldsymbol{a} + \boldsymbol{a}'\boldsymbol{b} + \boldsymbol{b}'\boldsymbol{a} + \boldsymbol{b}'\boldsymbol{b}$$

and also that $\boldsymbol{a}'\boldsymbol{b} = \boldsymbol{b}'\boldsymbol{a} = \sum_i b_i a_i$ (for vectors ONLY!).

Matrix and vector calculus

Let $\boldsymbol{x}$ and $\boldsymbol{a}$ be equal dimension vectors and $\boldsymbol{A}$ a matrix with column dimension the same as number of rows in $\boldsymbol{x}$. Then:

- $$\frac{d(\boldsymbol{x}'\boldsymbol{a})}{d\boldsymbol{x}} = \boldsymbol{a}$$
- $$\frac{d(\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x})}{d\boldsymbol{x}} = (\boldsymbol{A} + \boldsymbol{A}')\boldsymbol{x}$$
- i.e.
$$\boldsymbol{u}'\boldsymbol{u} = \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{y} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}$$

Solution

Returning to OLS

$$\frac{d(\boldsymbol{u}'\boldsymbol{u})}{d\boldsymbol{\beta}} = \frac{d(\boldsymbol{y}'\boldsymbol{y})}{d\boldsymbol{\beta}} - \frac{d(2\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{y})}{d\boldsymbol{\beta}} + \frac{d(\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta})}{d\boldsymbol{\beta}} =$$

$$= \mathbf{0} - 2\boldsymbol{X}'\boldsymbol{y} + 2\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}$$

First Order Condition:

$$\frac{d(\boldsymbol{u}'\boldsymbol{u})}{d\boldsymbol{\beta}} = \mathbf{0}$$

gives $\widehat{\boldsymbol{\beta}}_{OLS}$ which solves $\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}}_{OLS}$.
Left multiplication with $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ and

$$\widehat{\boldsymbol{\beta}}_{OLS} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$$

!!! The formula (and its derivation) is valid, regardless of dimensions of $\boldsymbol{X}$ and $\boldsymbol{\beta}$ (must match!)

!!! Second order condition is satisfied (but no proof here)

Alternative Solution

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u},$$

gives

$$\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{X}'\boldsymbol{X\beta} + \boldsymbol{X}'\boldsymbol{u}$$

which implies that:

$$\boldsymbol{\beta} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u},$$

and/or

$$(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u}.$$

i.e.

$$\widehat{\boldsymbol{\beta}}_{OLS} = \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u}$$

LS Assumptions

1. $\mathrm{E}(\boldsymbol{u}|\boldsymbol{X}) = \mathbf{0}$
2. $\boldsymbol{X}, \boldsymbol{Y}$ contain vectors of i.i.d. variables.
3. The 4th moments of $\boldsymbol{X}, \boldsymbol{Y}$ are finite.

Note that, e.g.

$$\mathrm{E}(\boldsymbol{u}|\boldsymbol{X}) = \mathbf{0} \rightarrow E\left(X'\boldsymbol{u}|\boldsymbol{X}\right) = 0$$

which implies that:

$$\begin{aligned} E\left(\widehat{\boldsymbol{\beta}}_{OLS}|\boldsymbol{X}\right) &= E\left(\boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u}|\boldsymbol{X}\right) \\ &= \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'E(\boldsymbol{u}|\boldsymbol{X}) = \boldsymbol{\beta} \end{aligned}$$

i.e. $\widehat{\boldsymbol{\beta}}_{OLS}$ is unbiased, since

$$E(\widehat{\boldsymbol{\beta}}_{OLS}) = E\left(E(\widehat{\boldsymbol{\beta}}_{OLS}|\boldsymbol{X})\right) = \boldsymbol{\beta}$$

more on this later

Properties

- Linearity
  $\widehat{\boldsymbol{\beta}}_{OLS} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{H}\boldsymbol{y}$, so that each element of $\widehat{\boldsymbol{\beta}}_{OLS}$ IS a linear combination of the observations in $\boldsymbol{y}$.
- Unbiasedness
- Covariance and consistency?: see later