Stata statistical software for data science

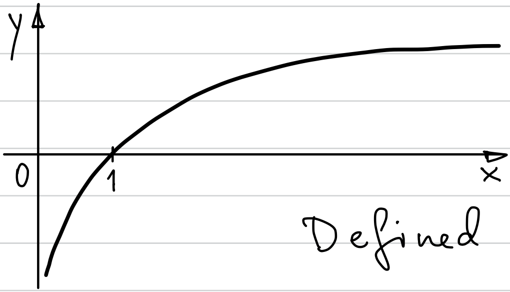
Key file formats:

- a) Data: .dta
- b) Programs: .do
- c) Graphs: .gph
- d) Logs: .smcl

Tutorial, Part 1

Ad 1c)

Logarithmic function:



Defined for $x > 0$.

Standardization: $x_{2i}^{s} = \frac{x_{2i} - \bar{x}_2}{sd(x_2)}$

Ad 2c)

$$D_t = \begin{cases} 1; & (spr_t \geq 0{,}8\, me(spr)) \overset{\text{"or"}}{\vee} (spr_t < \frac{2}{3}\overline{spr}) \\ 0; & \text{otherwise} \end{cases}$$

Ad 2d)

$\underline{spr_t}$

spr_{t-1}

spr_{t-4}

spr_{t+2}

$\Delta spr_t = spr_t - spr_{t-1}$

Basics of matrix algebra

- **Vector**, $a \in \mathbb{R}^{m\times 1}$, $b \in \mathbb{R}^{1\times n}$:

$$\underset{m\times 1}{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \qquad \underset{1\times n}{b} = [b_1 \dots b_n]$$

column vector

row vector

In econometrics, we have column vectors $\underset{n\times 1}{y}$, $\underset{n\times 1}{x_j}$, $\underset{n\times 1}{e}$, $\underset{k\times 1}{b}$, $\underset{k\times 1}{X^T y}$.

- **Matrix**, $A \in \mathbb{R}^{m\times n}$:

$$\underset{m\times n}{A} = \begin{array}{c} \left[\begin{array}{ccc} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{array}\right] \\ \begin{array}{ccc} a_1 & \qquad\quad & a_n \end{array} \end{array}$$

In econometrics, we have matrix $\underset{n\times k}{X}$.

- **Transpose**:

$$\underset{1\times m}{a^T} = [a_1 \dots a_m] \qquad \underset{n\times 1}{b^T} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \qquad \underset{n\times m}{A^T} = \begin{bmatrix} a_{11} & \dots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \dots & a_{mn} \end{bmatrix}$$

- **Square matrix**: $m = n$

$$(ABC)^T = C^T B^T A^T$$
$$(A^T)^T = A$$
$$(A^{-1})^T = (A^T)^{-1}$$

- **Symmetric square matrix**:

$A = A^T$ or $a_{ij} = a_{ji}, \forall i \neq j$

Example:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$$

diagonal

In econometrics, we have symmetric square matrices $\underset{k \times k}{X^TX}$ and $\underset{k \times k}{(X^TX)^{-1}}$.

- **Diagonal square matrix**:

$$A = \begin{bmatrix} a_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{bmatrix}$$

$a_{ij} = 0, \forall i \neq j$

- **Identity matrix**:

$$I = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}$$

- Inverse matrix:

Inverse matrix A^{-1} of a square matrix A exists, if and only if $A \cdot A^{-1} = A^{-1} \cdot A = I$.

For a matrix to be invertible, it has to be non-singular or regular: $\det(A) \neq 0$.

If an inverse matrix exists, it is the only inverse matrix.

Example 1:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$A^{-1} = \frac{1}{\det(A)} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}, \text{ where}$$

$$\det(A) = |A| = ad - bc.$$

Example 2:

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

$$A^{-1} = \frac{1}{\det(A)} \cdot \begin{bmatrix} c_{11} & -c_{12} & c_{13} \\ -c_{21} & c_{22} & -c_{23} \\ c_{31} & -c_{32} & c_{33} \end{bmatrix}^T$$

Rule of Sarrus:

$$\det(A) = \left|\begin{matrix} a & b & c \\ d & e & f \\ g & h & i \end{matrix}\right| \begin{matrix} a & b \\ d & e \\ g & h \end{matrix} =$$

$$= aei + bfg + cdh - ceg - afh - bdi$$

Cofactors:

$$c_{11} = \left|\begin{matrix} \cancel{a} & \cancel{b} & \cancel{c} \\ \cancel{d} & e & f \\ \cancel{g} & h & i \end{matrix}\right| = \left|\begin{matrix} e & f \\ h & i \end{matrix}\right| = ei - fh$$

$$c_{12} = \left|\begin{matrix} \cancel{a} & \cancel{b} & \cancel{c} \\ d & \cancel{e} & f \\ g & \cancel{h} & i \end{matrix}\right| = \left|\begin{matrix} d & f \\ g & i \end{matrix}\right| = di - fg$$

etc.