Indexing

Business Mathematics

VU University Amsterdam
Faculty of Economics and Business Administration

CONTENTS

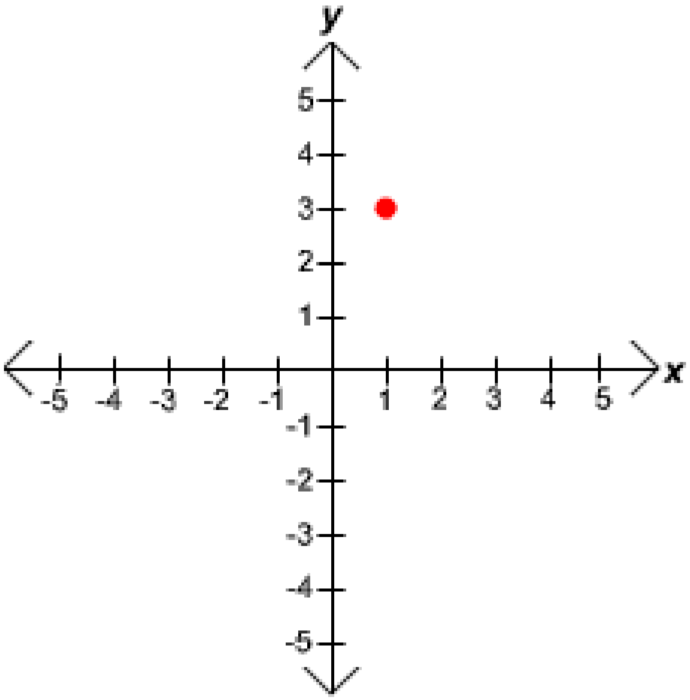
INDEXING VARIABLES

In the (x, y)-plane, we can identify each point with coordinates

- example: $(1,3)$
- or (x_1, x_2)

In a three-dimensional space, we use (x, y, z) or (x_1, x_2, x_3)

- and onward to $(x_1, x_2, x_3, x_4, \ldots)$



INDEXING VARIABLES

In a more abstract space $(p_1, q_1, p_2, q_2, \ldots)$

- where p_i is the price of product $i = 1,2, \ldots$
- and where q_i is the quantity of product $i = 1,2, \ldots$
- and where $i = 1,2, \ldots$ is an index variable (plural: indices)

Keep good track of the meaning and order of each coordinate

- $(1,3)$ is different from $(3,1)$

Mind that we use subscripts for the index, not superscripts

Observe that the index can be a variable as well

- compare x_1 with x_i

Usually (but not exclusively) index variable is i, j or k

VECTORS

It can be convenient to indicate a set of coordinates with one symbol

- e.g., $\mathbf{x} = (x_1, x_2, x_3)$

We refer to this as a vector

- to avoid confusion, also referred to as a row vector

Notation

- “normal” (scalar) variables are indicated by lowercase or uppercase italic print (x, a, π, Q, Λ)
- vectors are indicated by lowercase bold roman print ($\mathbf{x}$, $\mathbf{a}$, $\boldsymbol{\pi}$)
- handwritten (at your exam!) with an arrow ($\vec{x}$, $\vec{a}$, $\vec{\pi}$)

INDEXING OBSERVATIONS

We can also index different observations on the same variable

- example: $(p_1, p_2, p_3, \ldots)$ for price of house number 1, house number 2, etc.

Order of the elements less important, unless we have a second series for the same observations

- example: $(r_1, r_2, r_3, \ldots)$ for number of rooms of house number 1, house number 2, etc.
- or unless we have time series (1=1990, 2=1991, etc.)

Again, we can use vectors

- example: $\mathbf{p}$ and $\mathbf{r}$

DOUBLE INDEXING

We can combine indexing variables and observations

- example: $x_{3,2}$ is the number of rooms (variable 2) of house number 3 (observation 3)
- we can use variables for both indices ($x_{i,j}$)
- sometimes, the comma is skipped (x_{32}, x_{ij})

The order of the indices is crucial

- $x_{3,2}$ is very different from $x_{2,3}$

MATRICES

We can arrange double indexed variable in a “two-dimensional vector”

- e.g., $\mathbf{X} = \begin{pmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \\ x_{3,1} & x_{3,2} \end{pmatrix} = (\mathbf{x_1} \quad \mathbf{x_2})$

We refer to this as a <u>matrix</u>

Notation

- matrices are indicated by uppercase roman bold letters ($\mathbf{X}$, $\mathbf{A}$, $\mathbf{\Pi}$)
- handwritten just as a letter (X, A, Π), or with an arrow ($\vec{X}$, $\vec{A}$, $\vec{\Pi}$), or sometimes two arrows ($\vec{\vec{X}}$, $\vec{\vec{A}}$, $\vec{\vec{\Pi}}$)

Notice that the first index indicates the row and the second index denotes the column

MATRICES

This is an example of a data matrix

Each row represents an observation (country, household, company, etc.)

Each column represents a variable (price, employment, number of airplanes, etc.)

Each cell represents the variable for that observation, and can assume a numerical value

- example: $x_{i,j}$ is the value in row i and column j, referring to observation i on variable j

OLD EXAM QUESTION

22 October 2014, Q3c

A consultant expects that international debt might be correlated with national income. To investigate this, he collects international economic statistics in a data matrix **T**. See a small excerpt of **T** (with first column and first row indicating the meaning of each row and column) below.

Country	GDP/capita (USD/yr)	Inflation (%)	Debt/capita (USD)	Unemployment (%)
Afghanistan	678	5.6	92	35
Belarus	...	...	...	...

Develop a formula with the summation symbol for the covariance between GDP/capita and debt/capita, using the appropriately indexed elements of the matrix **T**. Introduce and define any symbols that you might need. (5 points)

OLD EXAM QUESTION

23 April 2015, Q2d

There exist urban legends on the relation between power failures and baby booms 9 months later. The government keeps a daily record of the length of the power interruption as a vector $\mathbf{p}' = (p_1, p_2, p_3, \dots)$. On most days it will be zero, but on some days it will be non-zero. The number of births is given by $\mathbf{b}' = (b_1, b_2, b_3, \dots)$. At this moment, both records have been kept for 3000 consecutive days. One way of studying the relation between $\mathbf{p}$ and $\mathbf{b}$ is by computing the correlation coefficient $r_{\mathbf{p},\mathbf{b}}$, but in a "delayed" form. Give an expression for the formula to use, taking into account the 9 months "delay". (5 points)

VU FEWEB