Advanced Quantitative Methods: Mathematics (p)review

Johan A. Elkink

University College Dublin

22 January 2013

Outline

Introduction

Topic: advanced quantitative methods in political science.

Introduction

Topic: advanced quantitative methods in political science.

Or, alternatively: basic econometrics, as applied in political science.

Introduction

Topic: advanced quantitative methods in political science.

Or, alternatively: basic econometrics, as applied in political science.

Or, alternatively: linear regression and some non-linear extensions.

Readings

- Peter Kennedy (2008), *A guide to econometrics*. 6th ed., Malden: Blackwell.
- Damodar N. Gujarati (2009), *Basic econometrics*. 5th ed., Boston: McGraw-Hill.
- Julian J. Faraway (2005), *Linear models with R*. Baco Raton: Chapman & Hill.

Readings

- Peter Kennedy (2008), *A guide to econometrics*. 6th ed., Malden: Blackwell.
- Damodar N. Gujarati (2009), *Basic econometrics*. 5th ed., Boston: McGraw-Hill.
- Julian J. Faraway (2005), *Linear models with R*. Baco Raton: Chapman & Hill.
- Andrew Gelman & Jennifer Hill (2007), *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.
- William H. Greene (2003), *Econometric analysis*. 5th ed., Upper Saddle River: Prentice Hall.

Topics

1	22/1	Mathematics review
2	29/1	Statistical estimators
3	5/2	Ordinary Least Squares
4	12/2	Hypothesis testing
5	19/2	Regression diagnostics
6	26/2	Time-series analysis
7	5/3	Causal inference
		Study break
8	26/3	Maximum Likelihood
9	2/4	Limited dependent variables I
10	9/4	Limited dependent variables II
11	16/4	Bootstrap and simulation
12	23/4	Multilevel data

Frequentist vs Bayesian

Frequentist statistics interprets **probability** as the frequency of occurrance in (hypothetically) many repetitions. E.g. if we throw this dice infinitely many times, what proportion of times would it be heads? We can here also talk of **conditional probabilities**: what would this frequency be if ... and some condition follows.

Frequentist vs Bayesian

Frequentist statistics interprets **probability** as the frequency of occurrance in (hypothetically) many repetitions. E.g. if we throw this dice infinitely many times, what proportion of times would it be heads? We can here also talk of **conditional probabilities**: what would this frequency be if ... and some condition follows.

Bayesian statistics interprets probability as a **belief**: if I throw this dice, what do you think is the chance of getting heads? We can now talk of conditional probabilities in a different way: how would your belief change given that ... and some condition follows.

Frequentist vs Bayesian

Frequentist statistics interprets **probability** as the frequency of occurrance in (hypothetically) many repetitions. E.g. if we throw this dice infinitely many times, what proportion of times would it be heads? We can here also talk of **conditional probabilities**: what would this frequency be if ... and some condition follows.

Bayesian statistics interprets probability as a **belief**: if I throw this dice, what do you think is the chance of getting heads? We can now talk of conditional probabilities in a different way: how would your belief change given that ... and some condition follows.

This course is a course in the frequentist analysis of regression models.

Homeworks

- 50 % Five homeworks
  due at specified dates
- 50 % Replication paper
  due May 5, 2012, 5 pm
- No exam

Homeworks

- 50 % Five homeworks
  due at specified dates
- 50 % Replication paper
  due May 5, 2012, 5 pm
- No exam

Working with others is a good idea

Grade conversions

Homeworks	UCD	TCD	Homeworks	UCD	TCD
95-100%	A+	A+	45-49%	E+	D
91-94%	A	A	37-44%	E	D
87-90%	A-	A	33-36%	E-	D
83-86%	B+	B+	0-32%	F	F
79-82%	B	B			
75-78%	B-	B			
70-74%	C+	C+			
66-69%	C	C			
62-65%	C	C			
58-61%	D+	C			
54-57%	D	C			
50-53%	D-	C			

Replication paper

- Find replicable paper *now* and check whether appropriate
- Contact authors asap if you need their data
- See Gary King, "Publication, publication"

Syllabus and website

- Website: http://www.joselkink.net/teaching
- Syllabus downloadable there
- Slides and notes on website
- Data for exercises on website
- Booklet with commands available

Outline

Vectors: examples

$$\mathbf{v} = \begin{bmatrix} 3 \\ 4 \\ 1 \\ 5 \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} 3.23 & 1.30 & 7.89 & 1.00 \end{bmatrix}$$

Vectors: examples

$$\mathbf{v} = \begin{bmatrix} 3 \\ 4 \\ 1 \\ 5 \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} 3.23 & 1.30 & 7.89 & 1.00 \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix}$$

Vectors: examples

$$\mathbf{v} = \begin{bmatrix} 3 \\ 4 \\ 1 \\ 5 \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} 3.23 & 1.30 & 7.89 & 1.00 \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix}$$

$\mathbf{v}$ and $\boldsymbol{\beta}$ are **column vectors**, while $\mathbf{w}$ is a **row vector**. When not specified, assume a column vector.

Vectors: transpose

$$\mathbf{v} = \begin{bmatrix} 3 \\ 4 \\ 1 \\ 5 \end{bmatrix} \qquad \mathbf{v}' = \begin{bmatrix} 3 & 4 & 1 & 5 \end{bmatrix}$$

Vectors: summation

$$\mathbf{v} = \begin{bmatrix} 3 \\ 5 \\ 9 \\ 1 \\ 3 \end{bmatrix}$$

$$\sum_{i}^{N} v_i = 3 + 5 + 9 + 1 + 3$$
$$= 21$$

Vectors: addition

$$\begin{bmatrix} 2 \\ 3 \\ 1 \\ 6 \end{bmatrix} + \begin{bmatrix} 6 \\ 3 \\ 9 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 6 \\ 10 \\ 8 \end{bmatrix}$$

Vectors: multiplication with scalar

$$\mathbf{v} = \begin{bmatrix} 2 \\ 3 \\ 1 \\ 6 \end{bmatrix}$$

$$3\mathbf{v} = \begin{bmatrix} 6 \\ 9 \\ 3 \\ 18 \end{bmatrix}$$

Vectors: inner product

$$\mathbf{v} = \begin{bmatrix} 5 \\ 3 \\ 1 \\ 3 \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} 7 \\ 6 \\ 8 \\ 1 \end{bmatrix}$$

$$\mathbf{v}'\mathbf{w} = 5 \cdot 7 + 3 \cdot 6 + 1 \cdot 8 + 3 \cdot 1 = 64$$

Vectors: outer product

$$\mathbf{v} = \begin{bmatrix} 5 \\ 3 \\ 1 \\ 3 \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} 7 \\ 6 \\ 8 \\ 1 \end{bmatrix}$$

$$\mathbf{v}\mathbf{w}' = \begin{bmatrix} 35 & 30 & 40 & 5 \\ 21 & 18 & 24 & 3 \\ 7 & 6 & 8 & 1 \\ 21 & 18 & 24 & 3 \end{bmatrix}$$

Matrices: examples

$$\mathbf{M} = \begin{bmatrix} 3 & 4 & 5 \\ 1 & 2 & 1 \\ 0 & 9 & 8 \end{bmatrix} \qquad \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Matrices: examples

$$\mathbf{M} = \begin{bmatrix} 3 & 4 & 5 \\ 1 & 2 & 1 \\ 0 & 9 & 8 \end{bmatrix} \qquad \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The latter is called an **identity matrix** and is a special type of **diagonal matrix**.

Both are **square matrices**.

Matrices: transpose

$$\mathbf{X} = \begin{bmatrix} 1 & 3 & 2 \\ 9 & 8 & 8 \\ 9 & 8 & 5 \end{bmatrix}$$

$$\mathbf{X}' = \begin{bmatrix} 1 & 9 & 9 \\ 3 & 8 & 8 \\ 2 & 8 & 5 \end{bmatrix}$$

Matrices: transpose

$$\mathbf{X} = \begin{bmatrix} 1 & 3 & 2 \\ 9 & 8 & 8 \\ 9 & 8 & 5 \end{bmatrix}$$

$$\mathbf{X}' = \begin{bmatrix} 1 & 9 & 9 \\ 3 & 8 & 8 \\ 2 & 8 & 5 \end{bmatrix}$$

$$(\mathbf{A}')' = \mathbf{A}$$

Matrices: indexing

$$\mathbf{X}_{4\times 3} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \end{bmatrix}$$

Matrices: indexing

$$\mathbf{X}_{4\times 3} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \end{bmatrix}$$

Always rows first, columns second.

Matrices: indexing

$$\mathbf{X}_{4\times 3} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \end{bmatrix}$$

Always rows first, columns second.

So $\mathbf{X}_{n\times k}$ is a matrix with n rows and k columns. $\mathbf{y}_{n\times 1}$ is a column vector with n elements.

Matrices: addition

$$\begin{bmatrix} 4 & 2 & 1 & 3 \\ 6 & 5 & 3 & 2 \\ 1\frac{1}{2} & 4 & 2\frac{1}{3} & 0 \end{bmatrix} + \begin{bmatrix} 7 & 8 & 2 & 0 \\ -3 & -2 & 3 & 1\frac{1}{2} \\ 4 & 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 11 & 10 & 3 & 3 \\ 3 & 3 & 6 & 3\frac{1}{2} \\ 5\frac{1}{2} & 6 & 3\frac{1}{3} & 1 \end{bmatrix}$$

Matrices: addition

$$\begin{bmatrix} 4 & 2 & 1 & 3 \\ 6 & 5 & 3 & 2 \\ 1\frac{1}{2} & 4 & 2\frac{1}{3} & 0 \end{bmatrix} + \begin{bmatrix} 7 & 8 & 2 & 0 \\ -3 & -2 & 3 & 1\frac{1}{2} \\ 4 & 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 11 & 10 & 3 & 3 \\ 3 & 3 & 6 & 3\frac{1}{2} \\ 5\frac{1}{2} & 6 & 3\frac{1}{3} & 1 \end{bmatrix}$$

$$(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$$

Matrices: multiplication with scalar

$$\mathbf{X} = \begin{bmatrix} 4 & 2 & 1 & 3 \\ 6 & 5 & 3 & 2 \\ 1\frac{1}{2} & 4 & 2\frac{1}{3} & 0 \end{bmatrix}$$

$$4\mathbf{X} = \begin{bmatrix} 16 & 8 & 4 & 12 \\ 24 & 20 & 12 & 8 \\ 6 & 16 & 7\frac{1}{3} & 0 \end{bmatrix}$$

Matrices: multiplication

$$\mathbf{A} = \begin{bmatrix} 6 & 4 \\ 1 & 3 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 3 & 2 & 3 \\ 4 & 2 & 1 \end{bmatrix}$$

$$\mathbf{AB} = \begin{bmatrix} 6 \cdot 3 + 4 \cdot 4 & 6 \cdot 2 + 4 \cdot 2 & 6 \cdot 3 + 4 \cdot 1 \\ 1 \cdot 3 + 3 \cdot 4 & 1 \cdot 2 + 3 \cdot 2 & 1 \cdot 3 + 3 \cdot 1 \end{bmatrix} = \begin{bmatrix} 34 & 20 & 22 \\ 15 & 8 & 6 \end{bmatrix}$$

Matrices: multiplication

$$\mathbf{A} = \begin{bmatrix} 6 & 4 \\ 1 & 3 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 3 & 2 & 3 \\ 4 & 2 & 1 \end{bmatrix}$$

$$\mathbf{AB} = \begin{bmatrix} 6 \cdot 3 + 4 \cdot 4 & 6 \cdot 2 + 4 \cdot 2 & 6 \cdot 3 + 4 \cdot 1 \\ 1 \cdot 3 + 3 \cdot 4 & 1 \cdot 2 + 3 \cdot 2 & 1 \cdot 3 + 3 \cdot 1 \end{bmatrix} = \begin{bmatrix} 34 & 20 & 22 \\ 15 & 8 & 6 \end{bmatrix}$$

$$(\mathbf{ABC})' = \mathbf{C}'\mathbf{B}'\mathbf{A}'$$
$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$$
$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$$
$$(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$$

Special matrices

Symmetric matrix	$\mathbf{A}' = \mathbf{A}$
Idempotent matrix	$\mathbf{A}^2 = \mathbf{A}$
Positive-definite matrix	$\mathbf{x}'\mathbf{A}\mathbf{x} > 0 \quad \forall \quad \mathbf{x} \neq 0$
Positive-semidefinite matrix	$\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0 \quad \forall \quad \mathbf{x} \neq 0$

Matrix rank

The **rank** of a matrix is the maximum number of independent columns or rows in the matrix. Columns of a matrix $\mathbf{X}$ are independent if for any $\mathbf{v} \neq 0$, $\mathbf{X}\mathbf{v} \neq 0$

Matrix rank

The **rank** of a matrix is the maximum number of independent columns or rows in the matrix. Columns of a matrix $\mathbf{X}$ are independent if for any $\mathbf{v} \neq 0$, $\mathbf{X}\mathbf{v} \neq 0$

$$r(\mathbf{A}) = r(\mathbf{A}') = r(\mathbf{A}'\mathbf{A}) = r(\mathbf{A}\mathbf{A}')$$
$$r(\mathbf{AB}) = min(r(\mathbf{A}), r(\mathbf{B}))$$

Matrix rank: example 1

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

For matrix $\mathbf{A}$, the three columns are independent and $r(\mathbf{A}) = 3$. There is no $\mathbf{v} \neq 0$ such that $\mathbf{Av} = 0$ (of course, if $\mathbf{v} = 0$, $\mathbf{Av} = \mathbf{A}0 = 0$).

Matrix rank: example 2

$$\mathbf{B} = \begin{bmatrix} 3 & 5 & 9 \\ 2 & 2 & 6 \\ 1 & 4 & 3 \end{bmatrix}$$

For matrix **B** the first and the last column vectors are linear combinations of each other: $\mathbf{b}_{\bullet 1} = \frac{1}{3}\mathbf{b}_{\bullet 3}$. At most two columns ($\mathbf{b}_{\bullet 1}$ and $\mathbf{b}_{\bullet 2}$ or $\mathbf{b}_{\bullet 2}$ and $\mathbf{b}_{\bullet 3}$) are independent, so $r(\mathbf{B}) = 2$. We could construct a matrix $\mathbf{v}$ such that $\mathbf{Bv} = 0$, namely $\mathbf{v} = \begin{bmatrix} 1 & 0 & -\frac{1}{3} \end{bmatrix}'$ (or any $\mathbf{v} = \begin{bmatrix} \alpha & 0 & -\frac{1}{3}\alpha \end{bmatrix}'$).

Matrix rank: example 3

$$\mathbf{C} = \begin{bmatrix} 3 & 5 & 11\frac{1}{2} \\ 2 & 2 & 7 \\ 1 & 4 & 5 \end{bmatrix}$$

$r(\mathbf{C}) = 2$. In this case, one cannot express one of the column vectors as a linear combination of another column vector, but one can express any of the column vectors as a linear combination of the two other column vectors. For example, $\mathbf{c}_{\bullet 3} = 3\mathbf{c}_{\bullet 1} + \frac{1}{2}\mathbf{c}_{\bullet 2}$ and thus when $\mathbf{v} = \begin{bmatrix} 3 & \frac{1}{2} & -1 \end{bmatrix}'$, $\mathbf{Cv} = 0$.

Matrix inverse

The **inverse** of a matrix is the matrix one would have to multiply with to get the identity matrix, i.e.:

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

Matrix inverse

The **inverse** of a matrix is the matrix one would have to multiply with to get the identity matrix, i.e.:

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

$$(\mathbf{A}^{-1})' = (\mathbf{A}')^{-1}$$
$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$

Matrices: trace

The **trace** of a matrix is the sum of the diagonal elements.

$$\mathbf{A} = \begin{bmatrix} 4 & 3 & 1 & 8 \\ 2 & 8 & 5 & 5 \\ 6 & 7 & 3 & 4 \\ 1 & 2 & 0 & 1 \end{bmatrix}$$

$$tr(\mathbf{A}) = sum(diag(\mathbf{A})) = 4 + 8 + 3 + 1 = 16$$

Outline

1. Course outline
2. Matrix algebra
3. Expectations and variances

Expected value

The **expected value** of a discrete random variable x is:

$$E(x) = \sum_{i}^{N} P(x_i) \cdot x_i,$$

whereby N is the total number of possible outcomes in S.

Expected value

The **expected value** of a discrete random variable x is:

$$E(x) = \sum_{i}^{N} P(x_i) \cdot x_i,$$

whereby N is the total number of possible outcomes in S.

The expected value is the **mean** (μ) of a random variable (not to be confused with the mean of a particular set of observations).

Moments

- moments are $E(x)^n$, $n = 1, 2, ...$
  - So μ_x is the first moment of x
- central moments are $E(x - E(x))^n$, $n = 1, 2, ...$
- absolute moments are $E(|x|)^n$, $n = 1, 2, ...$
- absolute central moments are $E(|x - E(x)|)^n$, $n = 1, 2, ...$

Variance

The second central moment is called the **variance**, so:

$$var(x) = E(x - E(x))^2 = E(x - \mu_x)^2$$
$$= \sum_{i}^{N} P(x_i) \cdot (x_i - E(x))^2$$

Variance

The second central moment is called the **variance**, so:

$$var(x) = E(x - E(x))^2 = E(x - \mu_x)^2$$
$$= \sum_{i}^{N} P(x_i) \cdot (x_i - E(x))^2$$

The **standard deviation** is the square root of the variance:

$$sd(x) = \sqrt{var(x)} = \sqrt{E(x - E(x))^2}$$

Covariance

The **covariance** of two random variables x and y is defined as:

$$cov(x, y) = E[(x - E(x))(y - E(y))]$$

Covariance

The **covariance** of two random variables x and y is defined as:

$$cov(x,y) = E[(x - E(x))(y - E(y))]$$

If x and y are **independent** of each other, $cov(x,y) = 0$.

Variance of a sum

$$var(x+y) = var(x) + var(y) + 2cov(x,y)$$
$$var(x-y) = var(x) + var(y) - 2cov(x,y)$$

Exercise

Let x be a random variable containing only ones and zeros. Let π be the probability of observing a one. Show that

1. $E(x) = \pi$
2. $var(x) = \pi(1 - \pi)$