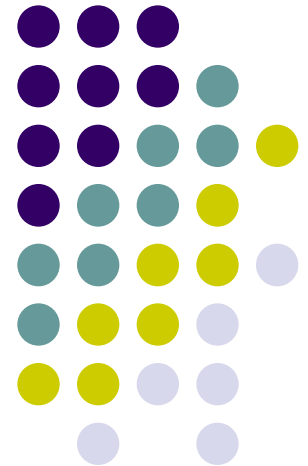


Polimetrics

Lecture 7 – Lab session
Let's run Wordscores!



How to run WORDSCORES

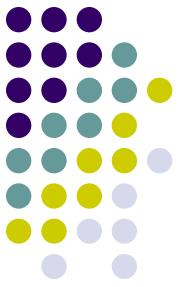


The Wordscores algorithm is implemented as a function in R

The estimation involves the following steps (quite similar to the WORDFISH ones..):

- (1) document processing (we already discussed about it!)
- (2) creation of a word count dataset (we already discussed about it! Remember the difference between Jfreq 2.5 and Jfreq 5.4!!!)
- (3) running Wordscores in R using the Austin package

Running WORDSCORES with R



We will employ the Austin package (that includes among the other things also Wordfish)

<http://conjugateprior.org/software/austin/>

Running WORDSCORES



There are three steps involved in running Wordscores:
loading the data, setting Wordscores options, and
running the code

1. Loading the data

If you have created the word count matrix outside of R, you
should load the word count data

Running WORDSCORES



2. Setting the options

Next, we need to tell Austin which manifestos are used as reference texts.

For this, we need to take note at which position the reference text are located. For example, if I want to use as reference texts the texts at position 4 and 3 in my dataset, I would write:

```
ref <- getdocs(data, c(4,3))
```

where *data* is my word count data

If the first two were the reference texts, for example, simply change this to

```
ref <- getdocs(data, c(1,2))
```

Running WORDSCORES



2. Setting the options

Next we need to assign scores to the reference texts. Here we use two reference texts, one with known issue position 3.2, and one with known issue position 19.7

```
ws <- classic.wordscores(ref, scores=c(19.7, 3.2))
```

Make sure that the order of reference texts is the same (i.e., for reference text 4 = 19.7, and for reference text 3 = 3.2 in the previous example)!!!

If you have additional reference texts, change the code into something like `scores=c(19.7, 3.2, 8.5)`

Running WORDSCORES



2. Setting the options

As an optional step, we can check the assignment using:

summary(ws)

To inspect the word scores you can type:

ws\$pi

Running WORDSCORES



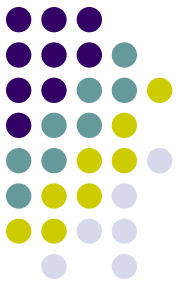
3. Running the code

So far we have set the reference texts. Now we want to predict the scores for the virgin texts. Type the following.

```
vir <- getdocs (data, c(1,2,5,6,7))
```

The text you identify as virgin texts here `c(1,2,5,6,7)` are complementary to the ones we identified as reference texts `c(4,3)`

Running WORDSCORES



3. Running the code

To get predictions of the positions of the virgin texts, we type the following.

```
scores <- predict(ws, newdata=vir)
```

You will now be presented the predicted scores for the virgin texts

The results include predicted scores, and rescaled scores that take into consideration the variance of the reference scores

Running WORDSCORES



WORDSCORES output

The estimation output can easily be called from the list object results for plotting purposes or further analysis.

The following output is available:

summary(scores) summary of the results

scores\$scores original scores of the documents

scores\$Std. Err. standard errors of the original scores

scores\$Rescaled rescaled scores of the documents

scores\$Lower lower 95% c.i. for the rescaled scores

scores\$Upper upper 95% c.i. for the rescaled scores

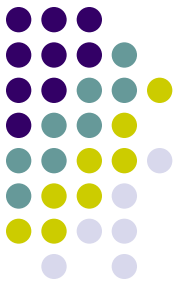
Wordscores in Stata? Yes you can!



Take a look at here!

http://www.tcd.ie/Political_Science/wordscores/software.html

Running WORDSCORES



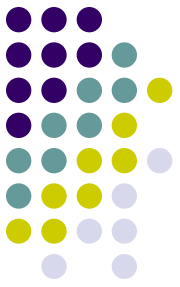
Let's see the UK 1992/1997 case with party manifesto case

First scenario: a) UK 1992 party manifestos as reference texts to estimate the UK 1997 party manifestos; b) estimate economic policy scores: Lab (UKLAB92a): 5.35; LibDem (UKLIBDEM92a): 8.21; Cons (UKCONS92a): 17.21

Second scenario: a) UK 1992 party manifestos as reference texts to estimate the UK 1997 party manifestos; b) social policy scores: Lab (UKLAB92a): 6.87; LibDem (UKLIBDEM92a): 6.53; Cons (UKCONS92a): 15.34

As encoding in Jfreq: UTF-8, no stemmer, no stop-words

Running WORDSCORES



Now run the analysis with respect to the Irish 1992/1997 party manifesto case under the following scenario:

- a) Irish 1992 party manifestos as reference texts to estimate the Irish 1997 party manifestos;
- b) estimate economic policy scores: DL (IREDL92) 4.50; Labour (IRELAB92): 6.88; Finn Fàil (IREFF92): 13.13; Fine Gael: 15 (IREFG92); PD (IREPD92): 17.63