

Computational Psycholinguistics, Day 6

LSA and LDA

Roger Levy

UC San Diego
Department of Linguistics

LSA Summer Institute
24 July 2007

The original word/document matrix

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

SVD with nine dimensions

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

MSE: 0

SVD with 8 dimensions

$$\begin{pmatrix}
 -0.01 & 0.04 & 0 & -0.01 & -0.04 & -0.09 & 0 & 0.08 & -0.07 \\
 0 & 0.01 & 0 & 0 & -0.01 & -0.02 & 0 & 0.02 & -0.02 \\
 0.01 & -0.04 & 0 & 0.02 & 0.05 & 0.11 & 0 & -0.09 & 0.08 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0.01 & -0.02 & 0 & 0.01 & 0.03 & 0.06 & 0 & -0.05 & 0.04 \\
 0 & 0 & 0 & 0 & -0.01 & -0.01 & 0 & 0.01 & -0.01 \\
 0 & 0 & 0 & 0 & -0.01 & -0.01 & 0 & 0.01 & -0.01 \\
 0 & 0.01 & 0 & -0.01 & -0.02 & -0.04 & 0 & 0.03 & -0.03 \\
 -0.01 & 0.05 & 0.01 & -0.02 & -0.06 & -0.13 & 0 & 0.11 & -0.1 \\
 -0.01 & 0.02 & 0 & -0.01 & -0.02 & -0.05 & 0 & 0.04 & -0.04 \\
 0.01 & -0.02 & 0 & 0.01 & 0.02 & 0.05 & 0 & -0.04 & 0.04 \\
 0 & -0.02 & 0 & 0.01 & 0.02 & 0.04 & 0 & -0.03 & 0.03
 \end{pmatrix}$$

MSE: 0.13

SVD with 7 dimensions

$$\begin{pmatrix} -0.01 & 0.04 & 0 & -0.01 & -0.04 & -0.08 & -0.02 & 0.09 & -0.07 \\ 0 & 0.01 & 0 & 0 & -0.01 & -0.02 & 0 & 0.02 & -0.02 \\ 0.01 & -0.05 & 0 & 0.02 & 0.05 & 0.1 & 0.02 & -0.11 & 0.08 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.01 & -0.02 & 0 & 0.01 & 0.03 & 0.05 & 0.01 & -0.06 & 0.05 \\ 0 & 0.01 & 0 & 0 & -0.01 & -0.01 & -0.01 & 0.01 & -0.01 \\ 0 & 0.01 & 0 & 0 & -0.01 & -0.01 & -0.01 & 0.01 & -0.01 \\ 0 & 0.02 & 0 & -0.01 & -0.02 & -0.03 & -0.01 & 0.04 & -0.03 \\ -0.01 & 0.05 & 0.01 & -0.02 & -0.06 & -0.12 & -0.01 & 0.12 & -0.1 \\ 0 & 0.01 & 0 & 0 & -0.01 & -0.12 & 0.11 & -0.02 & -0.03 \\ 0 & 0 & 0 & 0 & 0 & 0.23 & -0.29 & 0.13 & 0.01 \\ 0.01 & -0.03 & 0 & 0.01 & 0.04 & -0.13 & 0.29 & -0.21 & 0.06 \end{pmatrix}$$

MSE: 0.45

SVD with 6 dimensions

$$\begin{pmatrix} -0.09 & 0.23 & 0.11 & -0.13 & -0.34 & 0.07 & 0.04 & -0.02 & -0.09 \\ 0.01 & -0.02 & -0.01 & 0.01 & 0.03 & -0.04 & -0.01 & 0.04 & -0.02 \\ 0.06 & -0.16 & -0.07 & 0.08 & 0.22 & 0.01 & -0.02 & -0.05 & 0.09 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.03 & -0.09 & -0.04 & 0.05 & 0.12 & 0 & -0.01 & -0.02 & 0.05 \\ -0.04 & 0.11 & 0.06 & -0.07 & -0.17 & 0.07 & 0.03 & -0.05 & -0.02 \\ -0.04 & 0.11 & 0.06 & -0.07 & -0.17 & 0.07 & 0.03 & -0.05 & -0.02 \\ -0.01 & 0.03 & 0.01 & -0.01 & -0.04 & -0.02 & 0 & 0.03 & -0.03 \\ 0.06 & -0.12 & -0.09 & 0.09 & 0.21 & -0.26 & -0.07 & 0.22 & -0.08 \\ 0.04 & -0.09 & -0.06 & 0.06 & 0.15 & -0.2 & 0.07 & 0.04 & -0.02 \\ -0.02 & 0.06 & 0.03 & -0.04 & -0.09 & 0.27 & -0.27 & 0.09 & 0.01 \\ -0.04 & 0.09 & 0.06 & -0.06 & -0.15 & -0.03 & 0.33 & -0.28 & 0.05 \end{pmatrix}$$

MSE: 1.16

SVD with 5 dimensions

$$\begin{pmatrix} -0.12 & 0.11 & 0.43 & -0.29 & -0.32 & -0.06 & -0.05 & -0.02 & 0.08 \\ 0.06 & 0.15 & -0.48 & 0.25 & 0.01 & 0.15 & 0.12 & 0.04 & -0.25 \\ 0.03 & -0.24 & 0.18 & -0.04 & 0.23 & -0.09 & -0.09 & -0.05 & 0.21 \\ 0.04 & 0.13 & -0.36 & 0.19 & -0.02 & 0.15 & 0.11 & 0 & -0.18 \\ 0.01 & -0.15 & 0.16 & -0.05 & 0.13 & -0.08 & -0.07 & -0.02 & 0.15 \\ -0.06 & 0.05 & 0.22 & -0.15 & -0.16 & 0.01 & -0.02 & -0.05 & 0.06 \\ -0.06 & 0.05 & 0.22 & -0.15 & -0.16 & 0.01 & -0.02 & -0.05 & 0.06 \\ 0.02 & 0.12 & -0.25 & 0.12 & -0.05 & 0.08 & 0.07 & 0.03 & -0.16 \\ 0.06 & -0.09 & -0.16 & 0.13 & 0.21 & -0.22 & -0.05 & 0.22 & -0.12 \\ 0 & -0.22 & 0.31 & -0.13 & 0.17 & -0.35 & -0.04 & 0.04 & 0.17 \\ -0.01 & 0.1 & -0.08 & 0.02 & -0.1 & 0.32 & -0.24 & 0.09 & -0.05 \\ -0.01 & 0.18 & -0.2 & 0.07 & -0.17 & 0.08 & 0.41 & -0.28 & -0.09 \end{pmatrix}$$

MSE: 2.87

SVD with 4 dimensions

$$\begin{pmatrix} -0.12 & 0.08 & 0.49 & -0.33 & -0.27 & 0 & 0.01 & 0.01 & -0.02 \\ 0.04 & 0.24 & -0.64 & 0.34 & -0.13 & -0.02 & -0.02 & -0.03 & 0 \\ 0.04 & -0.27 & 0.24 & -0.07 & 0.29 & -0.03 & -0.03 & -0.02 & 0.12 \\ 0.02 & 0.23 & -0.55 & 0.29 & -0.18 & -0.04 & -0.07 & -0.08 & 0.12 \\ 0.02 & -0.2 & 0.25 & -0.1 & 0.21 & 0.02 & 0.02 & 0.01 & 0.01 \\ -0.07 & 0.08 & 0.17 & -0.12 & -0.2 & -0.04 & -0.06 & -0.06 & 0.13 \\ -0.07 & 0.08 & 0.17 & -0.12 & -0.2 & -0.04 & -0.06 & -0.06 & 0.13 \\ 0.01 & 0.15 & -0.31 & 0.15 & -0.1 & 0.02 & 0.01 & 0 & -0.05 \\ 0.1 & -0.26 & 0.14 & -0.04 & 0.47 & 0.09 & 0.23 & 0.34 & -0.61 \\ -0.04 & -0.04 & -0.02 & 0.06 & -0.12 & -0.71 & -0.35 & -0.09 & 0.71 \\ -0.01 & 0.07 & -0.04 & 0 & -0.06 & 0.36 & -0.2 & 0.11 & -0.11 \\ 0.01 & 0.09 & -0.03 & -0.02 & -0.02 & 0.25 & 0.57 & -0.21 & -0.36 \end{pmatrix}$$

MSE: 5.13

SVD with 3 dimensions

$$\begin{pmatrix} -0.76 & 0.06 & 0.52 & -0.14 & -0.17 & 0.01 & 0.02 & 0.01 & -0.04 \\ -0.82 & 0.21 & -0.61 & 0.58 & 0 & 0 & -0.01 & -0.02 & -0.02 \\ -0.89 & -0.3 & 0.28 & 0.19 & 0.43 & -0.01 & -0.01 & -0.01 & 0.09 \\ 0.17 & 0.24 & -0.56 & 0.24 & -0.21 & -0.04 & -0.07 & -0.08 & 0.12 \\ 0.54 & -0.19 & 0.23 & -0.25 & 0.13 & 0.01 & 0.01 & 0.01 & 0.02 \\ 0.05 & 0.08 & 0.17 & -0.15 & -0.22 & -0.04 & -0.06 & -0.06 & 0.14 \\ 0.05 & 0.08 & 0.17 & -0.15 & -0.22 & -0.04 & -0.06 & -0.06 & 0.14 \\ 0.3 & 0.16 & -0.33 & 0.07 & -0.15 & 0.01 & 0.01 & 0 & -0.05 \\ 0.05 & -0.26 & 0.14 & -0.03 & 0.48 & 0.1 & 0.23 & 0.34 & -0.61 \\ 0 & -0.04 & -0.03 & 0.04 & -0.13 & -0.71 & -0.35 & -0.1 & 0.71 \\ -0.01 & 0.07 & -0.04 & 0 & -0.06 & 0.36 & -0.2 & 0.11 & -0.11 \\ -0.01 & 0.09 & -0.03 & -0.02 & -0.02 & 0.25 & 0.57 & -0.21 & -0.36 \end{pmatrix}$$

MSE: 7.84

SVD with 2 dimensions

$$\begin{pmatrix} -0.84 & 0.4 & 0.38 & -0.53 & 0.18 & -0.05 & -0.12 & -0.16 & -0.09 \\ -0.86 & 0.37 & -0.67 & 0.4 & 0.16 & -0.03 & -0.07 & -0.1 & -0.04 \\ -0.85 & -0.49 & 0.36 & 0.41 & 0.24 & 0.02 & 0.06 & 0.09 & 0.12 \\ 0.26 & -0.16 & -0.39 & 0.7 & -0.61 & 0.03 & 0.08 & 0.12 & 0.19 \\ 0.45 & 0.23 & 0.05 & -0.73 & 0.56 & -0.07 & -0.15 & -0.21 & -0.05 \\ 0.16 & -0.42 & 0.38 & 0.42 & -0.72 & 0.06 & 0.13 & 0.19 & 0.22 \\ 0.16 & -0.42 & 0.38 & 0.42 & -0.72 & 0.06 & 0.13 & 0.19 & 0.22 \\ 0.22 & 0.55 & -0.49 & -0.37 & 0.24 & -0.07 & -0.14 & -0.2 & -0.11 \\ 0.1 & -0.47 & 0.23 & 0.21 & 0.27 & 0.14 & 0.31 & 0.44 & -0.58 \\ -0.06 & 0.23 & -0.14 & -0.27 & 0.14 & -0.76 & -0.45 & -0.23 & 0.66 \\ -0.06 & 0.34 & -0.15 & -0.3 & 0.2 & 0.31 & -0.31 & -0.02 & -0.15 \\ -0.04 & 0.25 & -0.1 & -0.21 & 0.15 & 0.22 & 0.5 & -0.29 & -0.38 \end{pmatrix}$$

MSE: 13.38

SVD with 1 dimension

$$\begin{pmatrix}
 -0.85 & 0.45 & 0.34 & -0.6 & 0.21 & 0 & 0.01 & 0.02 & 0.06 \\
 -0.87 & 0.4 & -0.69 & 0.36 & 0.18 & 0 & 0.01 & 0.02 & 0.05 \\
 -0.84 & -0.51 & 0.37 & 0.44 & 0.22 & 0 & 0.01 & 0.02 & 0.07 \\
 0.27 & -0.18 & -0.38 & 0.73 & -0.62 & 0.01 & 0.02 & 0.03 & 0.11 \\
 0.43 & 0.3 & 0 & -0.83 & 0.6 & 0.01 & 0.03 & 0.05 & 0.18 \\
 0.17 & -0.46 & 0.41 & 0.48 & -0.75 & 0 & 0.01 & 0.02 & 0.07 \\
 0.17 & -0.46 & 0.41 & 0.48 & -0.75 & 0 & 0.01 & 0.02 & 0.07 \\
 0.2 & 0.61 & -0.53 & -0.46 & 0.28 & 0 & 0.01 & 0.02 & 0.08 \\
 0.14 & -0.58 & 0.32 & 0.37 & 0.19 & 0 & 0.01 & 0.02 & -0.94 \\
 0.01 & 0.03 & 0.02 & 0.02 & 0.01 & -1 & -1 & -1 & 0 \\
 0.02 & 0.07 & 0.06 & 0.07 & 0.03 & 0 & -1 & -1 & -0.99 \\
 0.02 & 0.06 & 0.05 & 0.06 & 0.03 & 0 & 0 & -1 & -0.99
 \end{pmatrix}$$

MSE: 19.84

References