Neural Information Processing: 2012-2013
Assignment 1

School of Informatics, University of Edinburgh

Instructor: Mark van Rossum

28th February 2013

Remember that plagiarism is a university offence. Please read the policy at:

`http://www.inf.ed.ac.uk/admin/ITO/DivisionalGuidelinesPlagiarism.html`

Practical information

You should produce a digital document for your assignment answers (e.g. with lyx) and submit this electronically using the `submit` command on a DICE machine. The format is e.g.

```
submit
msc nip 1 nipasst1.pdf
```

You can check the status of your submissions with the `show_submissions` command. NOTE: postscript or pdf formats are acceptable, other formats are not. Make sure that the file you submit prints ok on the DICE system, in particular when you produced it on a non-Unix machine.

Late submissions

Late submissions will receive a zero mark. Only evidence for illness or other serious reasons can prevent this at the discretion of the instructor. See

`http://www.inf.ed.ac.uk/teaching/years/msc/courseguide09.html#exam`

Infomax vs PCA

In this assignment we analyse ways information transmission in linear networks. We assume that the two-dimensional input **u** is a Gaussian distributed signal with correlation matrix $\begin{pmatrix} 1 & q \\ q & 1 \end{pmatrix}$ with $q > 0$ and zero mean. The output of the network is given by $\mathbf{v} = W\mathbf{u} + \mathbf{n}$, where **n** is independent Gaussian noise with variance σ^2. The weight matrix $W = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix}$ is constrained such that $w_{k1}^2 + w_{k2}^2 = 1$ for $k = 1, 2$.

Question 1: We first analyse Linkers's approach to maximize information transmission between **u** and **v** (see lecture notes). Give an expression for $H(\mathbf{v})$. Express the determinant of the correlation matrix of **v** so that it contains two terms: one that is independent of σ and a σ-dependent part.

Question 2: Calculate the information in the limit $\sigma \to 0$, when $W = I$ (the identity matrix).

Question 3: Show that in the limit $\sigma \to \infty$, $W = 1/\sqrt{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ maximizes the information.

Question 4: Suppose that network instead performs PCA, projecting the input onto the (only) two principal components. What is W in that case? How much information does the PCA solution transmit in the $\sigma \to 0$ limit? Compare to the above results.

Question 5: Plot the mutual information as a function of σ^2 for $W = I$, $W = 1/\sqrt{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, and the PCA solution for $q = 1/2$.