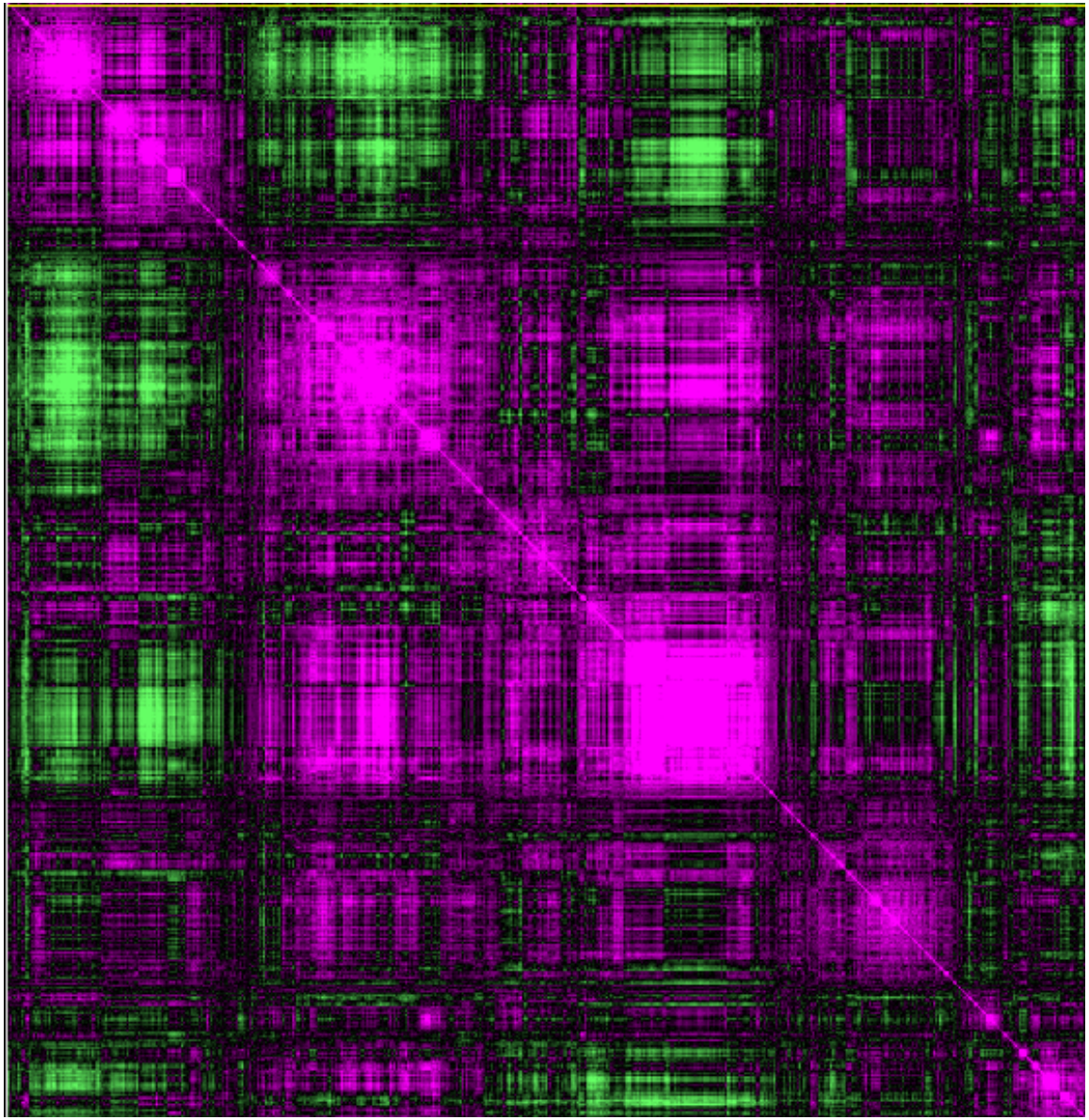
Practical Bioinformatics

Mark Voorhies

5/20/2019

Clustering exercises – Visualizing the distance matrix



Fun with logarithms

In log space, multiplication and division become addition and subtraction:

$$\begin{aligned} \log(xy) &= \log(x) + \log(y) \\ \log(x/y) &= \log(x) - \log(y) \end{aligned}$$

Fun with logarithms

In log space, multiplication and division become addition and subtraction:

$$\begin{aligned} \log(xy) &= \log(x) + \log(y) \\ \log(x/y) &= \log(x) - \log(y) \end{aligned}$$

Therefore, exponentiation becomes multiplication:

$$\log(x^y) = y \log(x)$$

Fun with logarithms

In log space, multiplication and division become addition and subtraction:

$$\begin{aligned} \log(xy) &= \log(x) + \log(y) \\ \log(x/y) &= \log(x) - \log(y) \end{aligned}$$

Therefore, exponentiation becomes multiplication:

$$\log(x^y) = y \log(x)$$

Also, we can change of the base of a logarithm like so:

$$\log_A(x) = \log(x)/\log(A)$$

Dictionaries

```
dictionary = {"A":"T","T":"A","G":"C","C":"G"}
dictionary["G"]
dictionary["N"] = "N"
dictionary.has_key("C")
```

Dictionaries

```
geneticCode = {"TTT":"F","TTC":"F","TTA":"L","TTG":"L",
               "CTT":"L","CTC":"L","CTA":"L","CTG":"L",
               "ATT":"I","ATC":"I","ATA":"I","ATG":"M",
               "GTT":"V","GTC":"V","GTA":"V","GTG":"V",

               "TCT":"S","TCC":"S","TCA":"S","TCG":"S",
               "CCT":"P","CCC":"P","CCA":"P","CCG":"P",
               "ACT":"T","ACC":"T","ACA":"T","ACG":"T",
               "GCT":"A","GCC":"A","GCA":"A","GCG":"A",

               "TAT":"Y","TAC":"Y","TAA":"*","TAG":"*",
               "CAT":"H","CAC":"H","CAA":"Q","CAG":"Q",
               "AAT":"N","AAC":"N","AAA":"K","AAG":"K",
               "GAT":"D","GAC":"D","GAA":"E","GAG":"E",

               "TGT":"C","TGC":"C","TGA":"*","TGG":"W",
               "CGT":"R","CGC":"R","CGA":"R","CGG":"R",
               "AGT":"S","AGC":"S","AGA":"R","AGG":"R",
               "GGT":"G","GGC":"G","GGA":"G","GGG":"G"}
```

Homework choice 1: Dictionaries

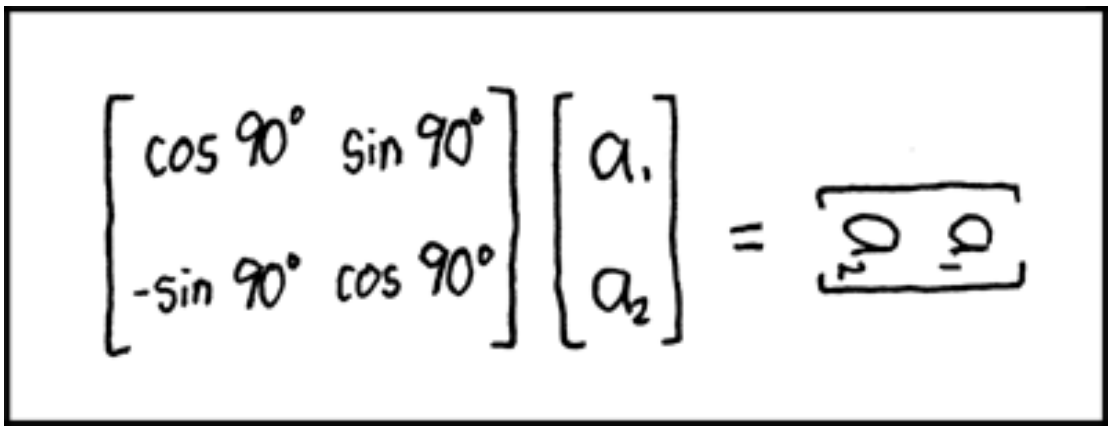
1. Write a function to return the antisense strand of a DNA sequence in 3'→5' orientation.
2. Write a function to return the complement of a DNA sequence in 5'→3' orientation.
3. Write a function to translate a DNA sequence

Homework choice 2: Linear Algebra

1. A two dimensional rotation, counterclockwise around the origin by θ, can be written as

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Try implementing rotations via numpy's dot function. *E.g.*, plot a set of two dimensional points, then rotate them and plot the new coordinates in a different color.



http://xkcd.com/184/