STA414 LEC0101 HW2

Solution in R

Prof Mark Ebden, January 31, 2018

A Python solution is available on Piazza. Below is an R solution:

```
library (mvtnorm) # A useful package for multivariate normals
X <- read.table("HW2.txt") # Load the 10x3 dataset
lnp <- numeric(3) # A 3-vector storing the log-likelihoods for cases a, b, and c
for (i in 1:5) {  # Loop over all 5x2 data points
  lnp[1] <- lnp[1] + log(dmvnorm(X[i,],c(3,2,2),diag(3))) + log(dmvnorm(X[i+5,],c(2,3,2),diag(3)))
  lnp[2] <- lnp[2] + log(dmvnorm(X[i,],c(2,3,2),diag(3))) + log(dmvnorm(X[i+5,],c(3,2,2),diag(3)))
  lnp[3] <- lnp[3] + log(dmvnorm(X[i,],c(2,2,3),diag(3))) + log(dmvnorm(X[i+5,],c(3,2,2),diag(3)))
}
print(rbind(exp(lnp),lnp)) # Print the likelihood and log-likelihood for each of the 3 cases
```

```
##               [,1]           [,2]           [,3]
##      2.279290e-163  2.371176e-128  8.934477e-146
## lnp  -3.744975e+02  -2.938675e+02  -3.339875e+02
```

Case (b), corresponding to `lnp[2]`, has the highest likelihood ($\sim 10^{-128}$) and therefore the answer is (b).

Note: if we didn't know which component created which point, the likelihood per point would be:

```
0.5*dmvnorm(X[i,],c(...),diag(3)) + 0.5*dmvnorm(X[i,],c(...),diag(3))
```

If you've coded a solution in an alternative language, please consider posting it on Piazza.