Jeffrey Kerley
Jakcqc@umsystem.edu
3/24/2022

MiniProject Four

##1.
Consider Gaussian density models in different dimensions.
(a) Write a program to find the maximum likelihood values ˆµ and ˆσ2.
Apply your program individually to each of the three features
xi of category ω1 in the table above.
we define our sample mean or maximum likelihood estimate as
= 1/n * sum k = 1 to n for (x sub k)
#our mle for ^σ2 is = 1/n * sum k = 1 to n for (x sub k - ^u)^2
#
our sample sigma or MLE for Sigma
= 1/ n-1 * sum k = 1 to n for
((x sub k - ^u sub n) * transpose(x sub k - ^u sub n))

```
> #mle mean x1
> w1MeanVec[1]
[1] -0.0709
> #mle mean x2
> w1MeanVec[2]
[1] -0.6047
> #mle mean x3
> w1MeanVec[3]
[1] -0.911

> #mle var2 x1
> w1VarVec[1]
[1] 0.9061773
> #mle var2 x2
> w1VarVec[2]
[1] 4.200715
> #mle var2 x3
> w1VarVec[3]
[1] 4.541949
```

(b) Modify your program to apply to two-dimensional Gaussian data p(x)~N(μ,Σ).
Apply your data to each of the three possible pairings of two features for ω1.

```
> # MLEmean for x1 and x2
> w1MeanVec[c(1,2)]
[1] -0.0709 -0.6047
> # MLEmean for x1 and x3
> w1MeanVec[c(1,3)]
[1] -0.0709 -0.9110
> # MLEmean for x2 and x3
> w1MeanVec[c(2,3)]
[1] -0.6047 -0.9110

> # MLESigma for x1 and x2
> w1Sigmax1x2 = mleSigma(w1[,c(1,2)],w1MeanVec[c(1,2)])
> w1Sigmax1x2
          [,1]      [,2]
[1,] 0.9061773 0.5677818
[2,] 0.5677818 4.2007148
> # MLESigma for x1 and x3
> w1Sigmax1x3 = mleSigma(w1[,c(1,3)],w1MeanVec[c(1,3)])
> w1Sigmax1x3
          [,1]      [,2]
[1,] 0.9061773 0.3940801
[2,] 0.3940801 4.5419490
> # MLESigma for x2 and x3
> w1Sigmax2x3 = mleSigma(w1[,c(2,3)],w1MeanVec[c(2,3)])
> w1Sigmax2x3
 [,1]      [,2]
[1,] 4.2007148 0.7337023
[2,] 0.7337023 4.5419490
```

(c) Modify your program to apply to three-dimensional Gaussian data.
Apply your data to the full three-dimensional data for ω1.

```
> #MLEmean x1 x2 x3
> w1MeanVec
        [,1]
[1,] -0.0709
```

[2,] -0.6047
[3,] -0.9110

```
> #MLEVar x1 x2 x3
> w1Sig
          [,1]      [,2]      [,3]
[1,] 0.9061773 0.5677818 0.3940801
[2,] 0.5677818 4.2007148 0.7337023
[3,] 0.3940801 0.7337023 4.5419490
```

(d) Assume your three-dimensional model is separable, so that
Σ = diag(σ21,σ22,σ23).
Write a program to estimate the mean and the diagonal components of Σ.
Apply your program to the data in ω2

```
> #MLe mean w2
> w2MeanVec
         [,1]
[1,] -0.11260
[2,]  0.42990
[3,]  0.00372
> #MLE var x1, x2, x3
> w2VarVec
            [,1]
[1,] 0.053925840
[2,] 0.045970090
[3,] 0.007265506
```

(e) Compare your results for the mean of each feature μi calculated
in the above ways. Explain why they are the same or different.

The means calculated in a) to c) are the "same" outside of dimensions. So each dimension will have the same mean no matter what, and with each added dimension our mean vector grows. When we look at the case for all uncorrelated dimensions, which is what the case d is saying, this will not impact our mle for variance squared and mean.

Since MLE relies upon assumptions about our variances and mean, assumptions about the correlation do not affect the outcome of MLE. Simply put, we use mle to find mean and variance, and then use our assumption to form our covariance matrix.

(f) Compare your results for the variance of each feature σ2i
calculated in the above ways. Explain why they are the same or different.

The variance for each feature stays the same for a) to b) as the relationship between each dimension is the same even as more dimensions are added. For d) as well, the variances for the diagonals are the same, but because of our assumption that our covariance is a diagonal, and that our data is separable, then we choose to ignore the covariance between each dimension. This gives us only the diagonal based on our assumption in d).

#problem 10
Suppose we know that the ten data points in category ω2
in the table above come from a three-dimensional uniform distribution
p(x|ω2)~ U(xl,xu).
a Suppose,however, that we do not have access to the
x3 components for the even-numbered data points.
#
(a) Write an EM program to estimate the six scalars comprising
xl and xu of the distribution. Start your estimate with
xl = (−2,−2,−2)t and x u= (+2,+2,+2)t.

xLow
[1] -0.400 0.054 -0.180
xHigh
[1] 0.38 0.69 0.12

(b) Compare your final esimate with that for the case when there is no missing data.

xLow no even x3 sample
[1] -0.400 0.054 -0.180
xHigh no even x3 sample
[1] 0.380 0.690 0.089

We can see how missing data would change the upper bounds for our uniform distribution. This method would simply be the MLE for each parameter which simply relies on the highest or lowest sample data.