MAD A4

Magnus Diamant (nrz659)

17. december 2019

Jeg har særligt samarbejdet med tre af mine studiekammerater: Ronnie Andersen (nqm290), William Lundsgaard Pedersen (wkb868) og Nicklas Christiansen (pwf334).

1 Opgave 1

1.1 a)

Jeg ved at min prior er betafordelt, og min likelihood er binomialfordelt. Dermed har posterior samme fordeling som prior, dvs. beta-fordeling (se side 103 om conjugate pairs).

Jeg kan nu bruge Bayes' formel som ser således ud:

$$P(r|y_N) = \frac{P(y_N|r)P(r)}{P(y_N)}$$

Jeg kan nu sætte de forskellige udtryk lig deres fordelinger ($P(y_N)$ fra nævneren er irrelevant da denne bare er konstant):

$P(y_N|r)$ er binomialfordelt og udregnes således:

$$P(y_N|r) = \binom{N}{k} r^k (1-r)^{N-k}$$

P(r) er betafordelt og er lig 1, når α og $\beta = 1$:

$$P(r) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} r^{\alpha-1}(1-r)^{\beta-1} = 1$$

For at finde posterior ganger jeg nu likelihood og prior sammen:

$$P(r)P(y_N|r) = 1 \cdot \binom{N}{k} r^k (1-r)^{N-k} = \binom{N}{k} r^k (1-r)^{N-k}$$

Jeg ved at dette vil være proportionalt med en betafordeling (fra 3.10 i bogen):

$$P(r)P(y_N|r) = \binom{N}{k} r^k (1-r)^{N-k} \propto \frac{\Gamma(\gamma+\delta)}{\Gamma(\gamma)\Gamma(\delta)} r^{\gamma-1}(1-r)^{\delta-1}$$

Fordi de to udtryk er proportionale med hinanden kan jeg fjerne konstanterne:

$$P(r)P(y_N|r) \propto r^k(1-r)^{N-k} \propto r^{\gamma-1}(1-r)^{\delta-1}$$

Det ses, at de to udtryk minder meget om hinanden. Jeg kan derfor isolere for γ og δ:

Først finder jeg γ

$$r^{\gamma-1} = r^k <=> \gamma - 1 = k <=> \gamma = k+1$$

Nu finder jeg δ:

$$(1-r)^{\delta-1} = (1-r)^{N-k} <=> \delta - 1 = N-k <=> \delta = N-k+1$$

Dermed har posterioren de to parametre $\gamma = k+1$ og $\delta = N-k+1$

1.2 b)

Dette er i bund og grund samme opgave som før, bortset fra at prior denne gang giver 2r. Fordelingerne er de samme. Jeg sætter derfor ind:

$$P(r)P(y_N|r) = 2r \cdot \binom{N}{k} r^k(1-r)^{N-k} = \binom{N}{k} r^k(1-r)^{N-k}$$

$$P(r)P(y_N|r) = 2r \cdot \binom{N}{k} r^k(1-r)^{N-k} \propto \frac{\Gamma(\gamma+\delta)}{\Gamma(\gamma)\Gamma(\delta)} r^{\gamma-1}(1-r)^{\delta-1}$$

Konstanterne fjernes igen:

$$P(r)P(y_N|r) = r \cdot r^k(1-r)^{N-k} \propto r^{\gamma-1}(1-r)^{\delta-1} <=>$$

$$P(r)P(y_N|r) = r^{k+1}(1-r)^{N-k} \propto r^{\gamma-1}(1-r)^{\delta-1}$$

Nu kan jeg isolere for γ og δ:

Først finder jeg γ:

$$r^{k+1} = r^{\gamma-1} <=> k+1 = \gamma - 1 <=> \gamma = k+2$$

Nu finder jeg δ:

$$(1-r)^{N-k} = (1-r)^{\delta-1} <=> N-k = \delta - 1 <=> \delta = N-k+1$$

Siden α og β var 1 i den forrige opgave og $\gamma = k+1$ og $\delta = N-k+1$, går jeg nu ud fra at $\alpha = 2$ og $\beta = 1$. Hvis man sætter dette ind i formlen går dette også op:

$$\frac{\Gamma(2+1)}{\Gamma(2)\Gamma(1)} r^{2-1}(1-r)^{1-1} = 2r$$

1.3 c)

Dette er i bund og grund samme opgave som før, bortset fra at prior denne gang giver $3r^2$. Fordelingerne er de samme. Jeg sætter derfor ind:

$$P(r)P(y_N|r) = 3r^2 \cdot \binom{N}{k} r^k (1-r)^{N-k} = \binom{N}{k} r^k (1-r)^{N-k}$$

$$P(r)P(y_N|r) = 3r^2 \cdot \binom{N}{k} r^k (1-r)^{N-k} \propto \frac{\Gamma(\gamma+\delta)}{\Gamma(\gamma)\Gamma(\delta)} r^{\gamma-1}(1-r)^{\delta-1}$$

Konstanterne fjernes igen:

$$P(r)P(y_N|r) = r^2 \cdot r^k (1-r)^{N-k} \propto r^{\gamma-1}(1-r)^{\delta-1} <=>$$

$$P(r)P(y_N|r) = r^{k+2}(1-r)^{N-k} \propto r^{\gamma-1}(1-r)^{\delta-1}$$

Nu kan jeg isolere for γ og δ:

Først finder jeg γ:

$$r^{k+2} = r^{\gamma-1} <=> k+2 = \gamma - 1 <=> \gamma = k+3$$

Nu finder jeg δ:

$$(1-r)^{N-k} = (1-r)^{\delta-1} <=> N-k = \delta - 1 <=> \delta = N-k+1$$

Siden α og β var 1 i den forrige opgave og $\gamma = k+1$ og $\delta = N-k+1$, går jeg nu ud fra at $\alpha = 3$ og $\beta = 1$. Hvis man sætter dette ind i formlen går dette også op:

$$\frac{\Gamma(3+1)}{\Gamma(3)\Gamma(1)} r^{3-1}(1-r)^{1-1} = 3r^2$$

2 Opgave 2

2.1 a)

Siden likelihooden er normalfordelt kan den defineres på følgende måde (fra s. 124 i bogen):

$$P(t|w, X, \sigma^2) = \mathcal{N}(X, w, \sigma^2 I_N) = \frac{1}{(2\pi)^{N/2}|\sigma^2 I_N|} e^{-\frac{1}{2}(x-\mu)^T(\sigma^2 I_N)^{-1}(x-\mu)} =$$

$$\frac{1}{(2\pi)^{N/2}|10 \cdot I_N|} e^{-\frac{1}{2}(x-\mu)^T(10 \cdot I_N)^{-1}(x-\mu)}$$

2.2 b)

Hvis både prior og likelihood er normalfordelt, må posterior også være det. Derfor kommer posterior til at se således ud:

$$P(w|t, X, \sigma^2) = \mathcal{N}(\mu_w, \Sigma_w)$$

, hvor $\Sigma_w = (\frac{1}{\sigma^2} X^T X + \Sigma_0^{-1})^{-1}$ og $\mu_w = \Sigma_w(\frac{1}{\sigma^2} X^T t + \Sigma_0^{-1}\mu_0)$

2.3 c) og d)

Jeg har desværre ikke haft tid til at løse disse opgaver.

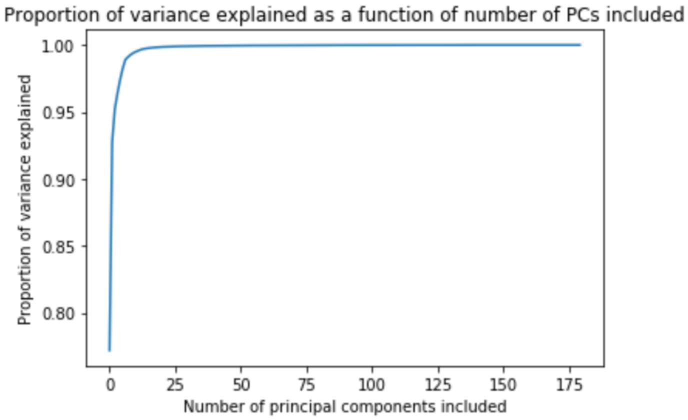
3 Opgave 3

3.1 a)

Se venligt nedenstående kode og resultater:

```
import numpy.matlib

def pca(data):
    data_cent = np.mean(data, 1)
    norm_diatoms = data - data_cent.reshape((-1,1))
    covariance = np.cov(norm_diatoms)

    PCevals, PCevecs = np.linalg.eigh(covariance)

    PCevals = PCevals[::-1]
    PCevecs = PCevecs[:,::-1]

    return PCevals, PCevecs, data_cent

PCevals, PCevecs, data_cent = pca(diatoms)
```

```
Proportion of variance explained by the first 1 principal components: 0.7718721493017529
Proportion of variance explained by the first 2 principal components: 0.9276996293043025
Proportion of variance explained by the first 3 principal components: 0.9521198453942007
Proportion of variance explained by the first 4 principal components: 0.9637878603999529
Proportion of variance explained by the first 5 principal components: 0.9739084497954094
Proportion of variance explained by the first 6 principal components: 0.98236065164916
Proportion of variance explained by the first 7 principal components: 0.9889975933245944
Proportion of variance explained by the first 8 principal components: 0.9910287023941854
Proportion of variance explained by the first 9 principal components: 0.9926692113360289
Proportion of variance explained by the first 10 principal components: 0.9939926229665051
```



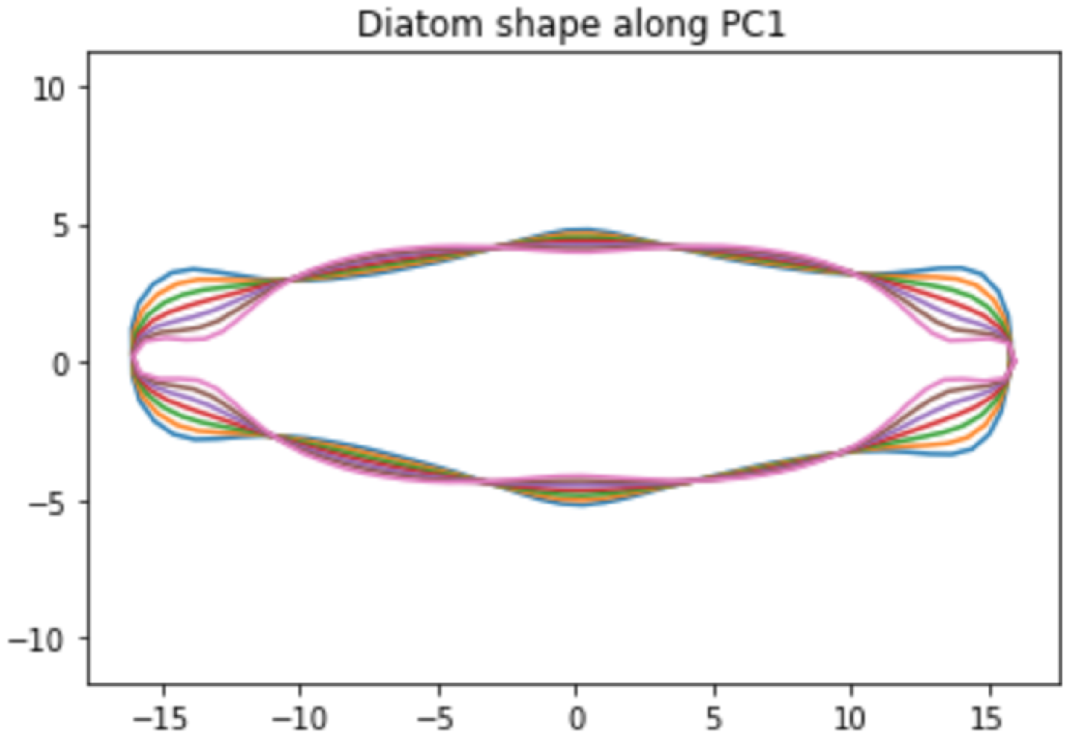
Det ses på resultaterne, at man kan genskabe vores datasæt med ca. 92% med de to første (og vigtigste) principal components. De tre første kan forklare lige over 95%, mens vi skal bruge 8 for at komme over 99%.

3.2 b)

Se venligst nedenstående kode og resultater:

```
e4 = PCevecs[:, 3]
lambda4 = PCevals[3]
std4 = np.sqrt(lambda4)

diatoms_along_pc = np.zeros((7, 180))
k = -3
for i in range(7):
    diatom = mean_diatom + e4 * std4 * k
    diatoms_along_pc[i] = diatom
    k = k+1

for i in range(7):
    plot_diatom(diatoms_along_pc[i])

plt.title('Diatom shape along PC1')
```

```
Text(0.5, 1.0, 'Diatom shape along PC1')
```



Det ses på plottet, den røde linje har en oval form, som ligner mean_diatom. Dette skyldes at den røde er plottet for PC nummer 4, som har en variance på 0 og derfor ligner mean_diatom.