CS 6190: Probabilistic Machine Learning Spring 2022

Homework 1

Handed out: 1 Feb, 2022
Due: 11:59pm, 18 Feb, 2022

- You are welcome to talk to other members of the class about the homework. I am more concerned that you understand the underlying concepts. However, you should write down your own solution. Please keep the class collaboration policy in mind.
- Feel free discuss the homework with the instructor or the TAs.
- Your written solutions should be brief and clear. You need to show your work, not just the final answer, but you do *not* need to write it in gory detail. Your assignment should be **no more than 10 pages**. Every extra page will cost a point.
- Handwritten solutions will not be accepted.
- The homework is due by **midnight of the due date**. Please submit the homework on Canvas.

Analytical problems [80 points + 30 bonus]

1. [8 points] A random vector, $\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}$ follows a multivariate Gaussian distribution,

$$p(\mathbf{x}) = \mathcal{N}\left(\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} | \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \right).$$

Show that the marginal distribution of $\mathbf{x}_1$ is $p(\mathbf{x}_1) = \mathcal{N}(\mathbf{x}_1|\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$.

Answer

We know that multivariate gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ is given by

$$p(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = |2\pi\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left(-(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right).$$

Writing out joint probability and substituting the matrices in the form described above, we get

$$p(\mathbf{x}_1, \mathbf{x}_2) \propto \exp\left(-\frac{1}{2} \begin{bmatrix} \mathbf{x}_1 - \mu_1 \\ \mathbf{x}_2 - \mu_2 \end{bmatrix}^T \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{x}_1 - \mu_1 \\ \mathbf{x}_2 - \mu_2 \end{bmatrix} \right).$$

Substituting the block matrix with block inversion formula and then multiplying out the outer matrices we have

$$p(x_1, x_2) = M \cdot N$$

where

$$M = \exp(-\frac{1}{2}(x_1 - \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2))^T(\Sigma_{11} - \Sigma_{21}^T\Sigma_{22}^{-1}\Sigma_{21})^{-1}(x_1 - \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)))$$

$$N = \exp(-\frac{1}{2}(x_2 - \mu_2)^T \Sigma_{22}^{-1}(x_2 - \mu_2)$$

We also know that for block matrices, $|M| = |M/H||H|$, and using this result to split normalisation term into two we get

$$\frac{1}{(2\pi^{(p+q)/2)}|\Sigma|^{1/2}} = \frac{1}{(2\pi^{(p+q)/2)}(|\Sigma/\Sigma_{11}||\Sigma_{11}|)^{1/2}} = \frac{1}{(2\pi)^{p/2}(|\Sigma/\Sigma_{11}|)^{1/2}} \cdot \frac{1}{(2\pi)^{q/2}(|\Sigma_{11}|)^{1/2}}$$

Doing this, we have just separated into products of marginal and conditional distribution. Integrating with respect to x_2, we get

$$p(\mathbf{x}_1) = (2\pi)^{q/2}(|\Sigma/\Sigma_{11}|)^{1/2} \exp\left(-\frac{1}{2}(\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1)\right).$$

hence proving the fact that $p(\mathbf{x}_1) = \mathcal{N}(\mathbf{x}_1|\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$.

2. [**Bonus**][10 points] Given a Gaussian random vector, $\mathbf{x} \sim \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$. We have a linear transformation, $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b} + \mathbf{z}$, where $\mathbf{A}$ and $\mathbf{b}$ are constants, $\mathbf{z}$ is another Gaussian random vector independent to $\mathbf{x}$, $p(\mathbf{z}) = \mathcal{N}(\mathbf{z}|\mathbf{0}, \boldsymbol{\Lambda})$. Show $\mathbf{y}$ follows Gaussian distribution as well, and derive its form. Hint: using characteristic function. You need to check the materials by yourself.

3. [8 points] Show the differential entropy of the a multivariate Gaussian distribution $\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is

$$\mathrm{H}[\mathbf{x}] = \frac{1}{2}\log|\boldsymbol{\Sigma}| + \frac{d}{2}(1 + \log 2\pi)$$

where d is the dimension of $\mathbf{x}$.

Answer

$$H(P) = -\int P(x)\log P(x)$$

$$H(P) = \int P(x)\left[\frac{D}{2}\log 2\pi + \frac{1}{2}\log|\Sigma| + \frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right]$$

Taking out the constant terms and writing the leftover integral in terms of expectation and trace we get

$$H(P) = \frac{D}{2}\log 2\pi + \frac{1}{2}\log|\Sigma| + \frac{1}{2}\mathbb{E}_P\big[\mathrm{Tr}(\Sigma^{-1}(x-\mu)(x-\mu)^T)\big]$$

Taking trace out and sending Expectation inside and rewriting we get

$$H(P) = \frac{D}{2}\log 2\pi + \frac{1}{2}\log|\Sigma| + \frac{1}{2}\mathrm{Tr}\big[\Sigma^{-1}\Sigma\big]$$

$$H(P) = \frac{D}{2}\log 2\pi + \frac{1}{2}\log|\Sigma| + \frac{1}{2}\mathrm{Tr}\big[I\big]$$

$$H(P) = \frac{D}{2}\log 2\pi + \frac{1}{2}\log|\Sigma| + \frac{D}{2}$$

4. [8 points] Derive the Kullback-Leibler divergence between two Gaussian distributions, $p(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $q(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\mathbf{m}, \Lambda)$, i.e., $\mathrm{KL}(q||p)$.

Answer

$$\mathrm{KL}(q||p) = \int q(x)\log\left(\frac{q(x)}{p(x)}\right) = \mathbb{E}\big[\log(q) - \log(p)\big]$$

As

$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = |2\pi|^{-k/2}|\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

$$\mathrm{KL}(q||p) = \frac{1}{2}\log\left(\frac{|\Sigma|}{|\Lambda|}\right) - \frac{1}{2}\mathbb{E}_q[(x-m)^T\Lambda^{-1}(x-m)] + \frac{1}{2}\mathbb{E}_q[(x-\mu)^T\Sigma^{-1}(x-\mu)]$$

Using the property of trace and the fact that 2nd and 3nd terms have scalar inside we can write

$$\mathrm{KL}(q||p) = \frac{1}{2}\log\left(\frac{|\Sigma|}{|\Lambda|}\right) - \frac{1}{2}\mathrm{Tr}[\Lambda\Lambda^{-1}] + (m-\mu)^T\Sigma^{-1}(m-\mu) + \mathrm{Tr}[\Sigma^{-1}\Sigma]$$

5. [8 points] Given a distribution in the exponential family,

$$p(\mathbf{x}|\boldsymbol{\eta}) = \frac{1}{Z(\boldsymbol{\eta})}h(\mathbf{x})\exp\big(-\mathbf{u}(\mathbf{x})^\top\boldsymbol{\eta}\big).$$

Show that

$$\frac{\partial^2 \log Z(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}^2} = \mathrm{cov}(\mathbf{u}(\mathbf{x})),$$

where cov is the covariance matrix.

Answer

Let $A(\boldsymbol{\eta}) = \log Z(\boldsymbol{\eta})$, then

$$\frac{\partial^2 A(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}^2} = \int \mathbf{u}(x)\exp(\mathbf{u}(x)^T\boldsymbol{\eta} - A(\boldsymbol{\eta}))(\mathbf{u}(x) - \frac{\partial A}{\partial \boldsymbol{\eta}})h(\mathbf{x})dx$$

$$\frac{\partial^2 A(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}^2} = \int \mathbf{u}^2(x)\frac{h(x)\exp(\mathbf{u}(x)^T\boldsymbol{\eta})}{Z(\boldsymbol{\eta})}dx - \frac{\partial A}{\partial \boldsymbol{\eta}}\int \mathbf{u}(x)\frac{h(x)\exp(\mathbf{u}(x)^T\boldsymbol{\eta})}{Z(\boldsymbol{\eta})}dx$$

$$\frac{\partial^2 A(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}^2} = \int \mathbf{u}^2(x)P(\mathbf{x}|\boldsymbol{\eta})dx - \frac{\partial A}{\partial \boldsymbol{\eta}}\int \mathbf{u}(x)P(\mathbf{x}|\boldsymbol{\eta})dx$$

$$\frac{\partial^2 A(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}^2} = \mathbb{E}(u^2(x)) - [\mathbb{E}(u(x))]^2 = \mathrm{cov}(\mathbf{u}(\mathbf{x}))$$

6. [4 points] Is $\log Z(\boldsymbol{\eta})$ convex or nonconvex? Why?

Answer

Let $A = \log Z(\boldsymbol{\eta})$ and $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2$ be such that $\boldsymbol{\eta} = \lambda\boldsymbol{\eta}_1 + (1-\lambda)\boldsymbol{\eta}_2$.

$$Z(\boldsymbol{\eta}) = \int \exp(u(x)^T(\lambda\boldsymbol{\eta}_1 + (1-\lambda)\boldsymbol{\eta}_2))h(x)dx$$

$$Z(\boldsymbol{\eta}) \le \left(\int \exp(u(x)^T\lambda\boldsymbol{\eta}_1)\frac{h(x)}{\lambda}dx\right)^\lambda \left(\int \exp(u(x)^T\lambda\boldsymbol{\eta}_2)\frac{h(x)}{1-\lambda}dx\right)^{(1-\lambda)}$$

$$Z(\boldsymbol{\eta}) \le \left(\int \exp(u(x)^T\boldsymbol{\eta}_1)\frac{h(x)}{\lambda}dx\right)^\lambda \left(\int \exp(u(x)^T\boldsymbol{\eta}_2)\frac{h(x)}{1-\lambda}dx\right)^{(1-\lambda)}$$

Taking log on both sides we can see that

$$A(\lambda\boldsymbol{\eta}_1 + (1-\lambda)\boldsymbol{\eta}) \le \lambda A(\boldsymbol{\eta}_1) + (1-\lambda)A(\boldsymbol{\eta}_2)$$

Using holders' inequality we have proved that $\log Z(\boldsymbol{\eta})$ is convex.

7. [8 points] Given two random variables $\mathbf{x}$ and $\mathbf{y}$, show that

$$I(\mathbf{x}, \mathbf{y}) = H[\mathbf{x}] - H[\mathbf{x}|\mathbf{y}]$$

where $I(\cdot, \cdot)$ is the mutual information and $H[\cdot]$ the entropy.

Answer

$$I(\mathbf{x}, \mathbf{y}) = \int\int P(\mathbf{x}, \mathbf{y}) \log \left(\frac{P(\mathbf{x}, \mathbf{y})}{P(\mathbf{x})P(\mathbf{y})} \right) d\mathbf{x} d\mathbf{y}$$

$$I(\mathbf{x}, \mathbf{y}) = \int\int P(\mathbf{x}, \mathbf{y}) \log \left(\frac{P(\mathbf{x}|\mathbf{y})}{P(\mathbf{x})} \right) d\mathbf{x} d\mathbf{y}$$

$$I(\mathbf{x}, \mathbf{y}) = -\int\int P(\mathbf{x}, \mathbf{y}) \log P(\mathbf{x})) d\mathbf{x} d\mathbf{y} + \int\int P(\mathbf{x}, \mathbf{y}) \log P(\mathbf{x}|\mathbf{y}) d\mathbf{x} d\mathbf{y}$$

$$I(\mathbf{x}, \mathbf{y}) = H(\mathbf{x}) - H(\mathbf{x}|\mathbf{y})$$

8. [24 points] Convert the following distributions into the form of the exponential-family distribution. Please give the mapping from the expectation parameters to the natural parameters, and also represent the log normalizer as a function of the natural parameters.

- Dirichlet distribution

  Answer

$$P(x|\alpha) = \frac{\Gamma(\sum_k \alpha_k)}{\prod_k \Gamma(_k)} \prod_k x_k^{\alpha_k - 1}$$

$$P(x|\alpha) = \exp\left(\log\Gamma(\sum_k \alpha_k) - \sum_k \log\Gamma(\alpha_k) + \sum_k (\alpha_k - 1)\log x_k \right)$$

$$P(x|\alpha) = \exp\left(\sum_k (\alpha_k - 1)\log x_k - (\sum_k \log\Gamma(\alpha_k) - \log\Gamma(\sum_k \alpha_k) \right)$$

  From above form we can see that $\boldsymbol{\eta} = \alpha - 1$, $A(\boldsymbol{\eta}) = \sum_k \log\Gamma(\alpha_k) - \log\Gamma(\sum_k \alpha_k)$, $u(x) = \log x$ and $h(x) = 1$

- Gamma distribution

  Answer

$$P(x|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x)$$

$$P(x|\alpha, \beta) = \exp(\alpha \log\beta + (\alpha - 1)\log x - \log\Gamma(\alpha) - \beta x)$$

$$P(x|\alpha, \beta) = \exp(-\beta x) \exp((\alpha - 1)\log x - (\log\Gamma(\alpha) - \alpha\log\beta))$$

  From above form we can see that $h(x) = exp(-\beta x)$, $u(x) = \log \mathbf{x}$, $A(\boldsymbol{\eta}) = \log\Gamma(\alpha) - \alpha\log\beta$ and $\boldsymbol{\eta} = \alpha - 1$

- Wishart distribution

  Answer

$$f_{\mathbf{X}}(\mathbf{X}, \mathbf{V}, n) = \frac{1}{2^{np/2}|\mathbf{V}|^{n/2}\Gamma_p(n/2)} |\mathbf{x}|^{(n-p-1)/2} \exp(-\frac{1}{2}\mathrm{tr}(\mathbf{V}^{-1}\mathbf{x}))$$

  We know that $\mathrm{tr}(A^T B) = vec(A^T) \cdot vec(B)$

$$f_{\mathbf{X}}(\mathbf{X}, \mathbf{V}, n) = \frac{1}{2^{np/2}|\mathbf{V}|^{n/2}\Gamma_p(n/2)} \exp((vec(\mathbf{X}), \log|\mathbf{x}|)^T(-\frac{vec(\mathbf{V}^{-1})}{2}, \frac{n-p-1}{2}))$$

  Comparing the form above against the exponential family form we get, $(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2) = (-\frac{vec(\mathbf{V}^{-1})}{2}, \frac{n-p-1}{2})$, $h(x) = 1$, $A(\boldsymbol{\eta}) = -(\boldsymbol{\eta}_2 + \frac{p+1}{2})\log|-\boldsymbol{\eta}_1| + \log\Gamma_p(\boldsymbol{\eta}_2 + \frac{p+1}{2})$, $T(x) = (vec(\mathbf{X}), \log|\mathbf{X}|)$

9. [6 points] Does student t distribution (including both the scalar and vector cases) belong to the exponential family? Why?

Answer

No, students' t distribution does not belong to the exponential family. For a distribution to belong to exponential family, there has to be a certain form where it can be represented as $\frac{1}{Z(\boldsymbol{\eta})} h(x) \exp(u(x)^T \boldsymbol{\eta})$. If we write out students t distribution

$$st(x|\mu, \lambda, \gamma) = \frac{\Gamma(\gamma/2 + 1/2)}{\Gamma(\gamma/2)} \left(\frac{\lambda}{\pi\gamma}\right)^{1/2} \left[1 + \frac{\lambda(x-\mu)^2}{\gamma}\right]^{-\gamma/2 - 1/2}$$

there's no way to get this into the representative form of exponential family of distribution.

10. [6 points] Does the mixture of Gaussian distribution belong to the exponential family? Why?

$$p(\mathbf{x}) = \frac{1}{2}\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) + \frac{1}{2}\mathcal{N}(\mathbf{x}|\mathbf{m}, \boldsymbol{\Lambda})$$

Answer

No the mixture of two gaussians doesn't belong to exponential family. The product of two gaussains will belong to exponential family because the terms in exponentials will get added and give us the desired form however that's not possible when two gaussians are added.

$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) + \mathcal{N}(\mathbf{x}|\mathbf{m}, \boldsymbol{\Lambda}) = |2\pi|^{-k/2} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right) +$$
$$|2\pi|^{-k/2} |\boldsymbol{\Lambda}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{m})^\top \boldsymbol{\Lambda}^{-1}(\mathbf{x} - \mathbf{m})\right)$$

11. [**Bonus**][20 points] Given a distribution in the exponential family $p(\mathbf{x}|\boldsymbol{\eta})$, where $\boldsymbol{\eta}$ are the natural parameters. As we discussed in the class, the distributions in the exponential family are often parameterized by their expectations, namely $\boldsymbol{\theta} = \mathbb{E}(\mathbf{u}(\mathbf{x}))$ where $\mathbf{u}(\mathbf{x})$ are the sufficient statistics (recall Gaussian and Bernoulli distributions). Given an arbitrary distribution $p(\mathbf{x}|\boldsymbol{\alpha})$, the Fisher information matrix in terms of the distribution parameters $\boldsymbol{\alpha}$ is defined as $\mathbf{F}(\boldsymbol{\alpha}) = \mathbb{E}_{p(\mathbf{x}|\boldsymbol{\alpha})}[-\frac{\partial^2 \log(p(\mathbf{x}|\boldsymbol{\alpha}))}{\partial \boldsymbol{\alpha}^2}]$.

    (a) [5 points] Show that if we calculate the Fisher Information matrix in terms of the natural parameters, we have $\mathbf{F}(\boldsymbol{\eta}) = \text{cov}(\mathbf{u}(\mathbf{x}))$.

    (b) [5 points] Show that $\frac{\partial \boldsymbol{\theta}}{\partial \boldsymbol{\eta}} = \mathbf{F}(\boldsymbol{\eta})$.

    (c) [10 points] Show that the Fisher information matrix in terms of the expectation parameters is the inverse of that in terms of the natural parameters, $\mathbf{F}(\boldsymbol{\theta}) = \mathbf{F}^{-1}(\boldsymbol{\eta})$.

    (d) [5 points] Suppose we observed dataset $\mathcal{D}$. Show that

    $$\frac{\partial \log p(\mathcal{D}|\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} \mathbf{F}(\boldsymbol{\eta})^{-1} = \frac{\partial \log p(\mathcal{D}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$$

    and

    $$\frac{\partial \log p(\mathcal{D}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \mathbf{F}(\boldsymbol{\theta})^{-1} = \frac{\partial \log p(\mathcal{D}|\boldsymbol{\eta})}{\partial \boldsymbol{\eta}}.$$

    Note that I choose the orientation of the gradient vector to be consistent with Jacobian. So, in this case, the gradient vector is a row vector (rather than a column vector). If you want to use a column vector to represent the gradient, you can move the information matrix to the left. It does not influence the conclusion.

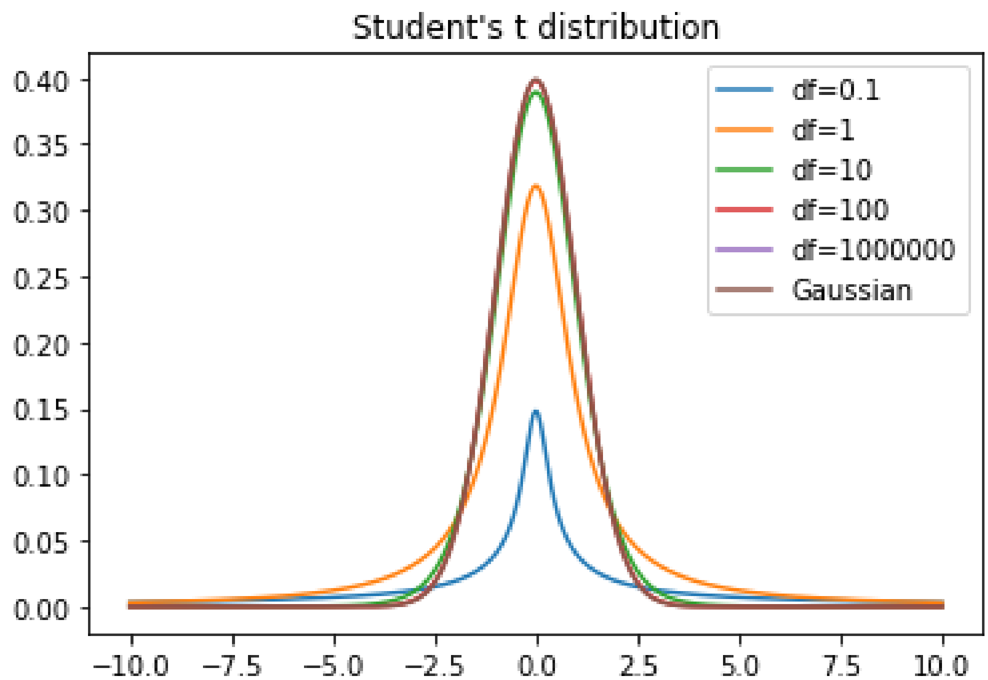
1 Practice [20 points]

1. [5 Points] Look into the student t's distribution. Let us set the mean and precision to be $\mu = 0$ and $\lambda = 1$. Vary the degree of freedom $\nu = 0.1, 1, 10, 100, 10^6$ and draw the density of the student t's distribution. Also, draw the density of the standard Gaussian distribution $\mathcal{N}(0, 1)$. Please place all the density curves in one figure. Show the legend. What can you observe?

   Answer

   Code can be found here

   Here's the plot showing all the different distributions.

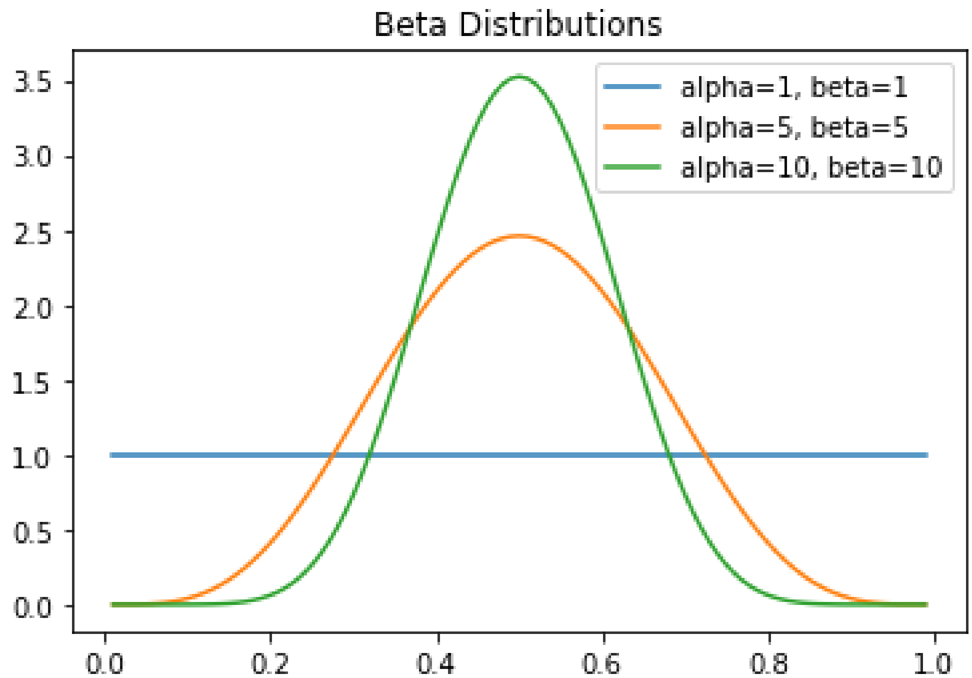
   

   From the plot above we can observe that for high degrees of freedom, student's t distribution becomes flat and almost converges into gaussian like distribution when degree of freedom is very small.

2. [5 points] Draw the density plots for Beta distributions: Beta(1,1), Beta(5, 5) and Beta (10, 10). Put the three density curves in one figure. What do you observe? Next draw the density plots for Beta(1, 2), Beta(5,6) and Beta(10, 11). Put the three density curves in another figure. What do you observe?

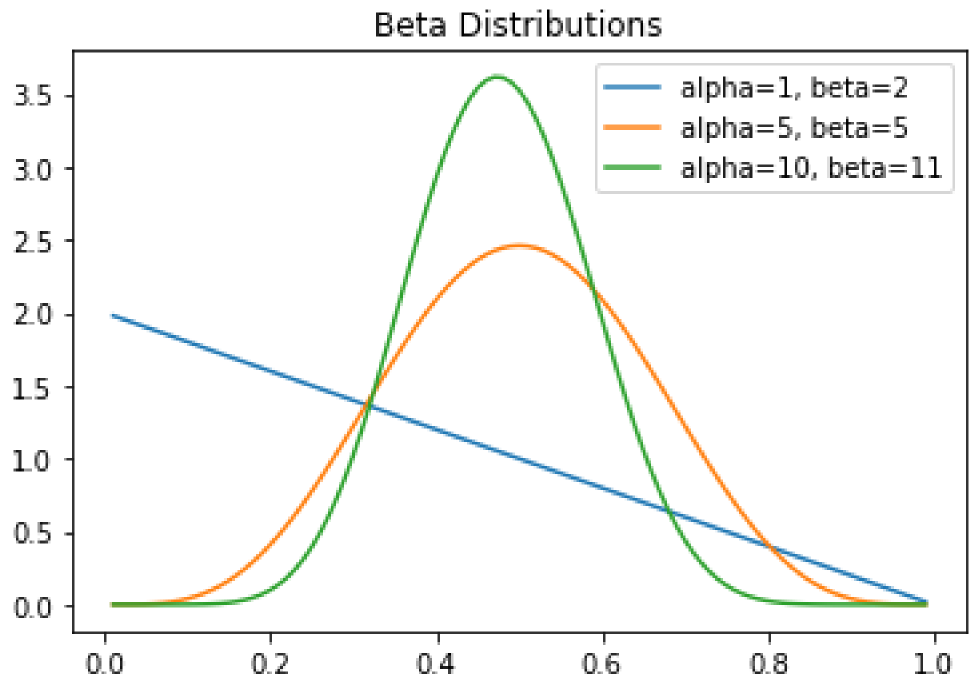
   Answer

   Here's the first plot.



From the plot above we can see that as value of α and β decreases, the curve becomes flatter and flatter.

From the 2nd plot (shown below), we can see that distributions become lopsided but still lower values of parameters indicate flatter distributions.

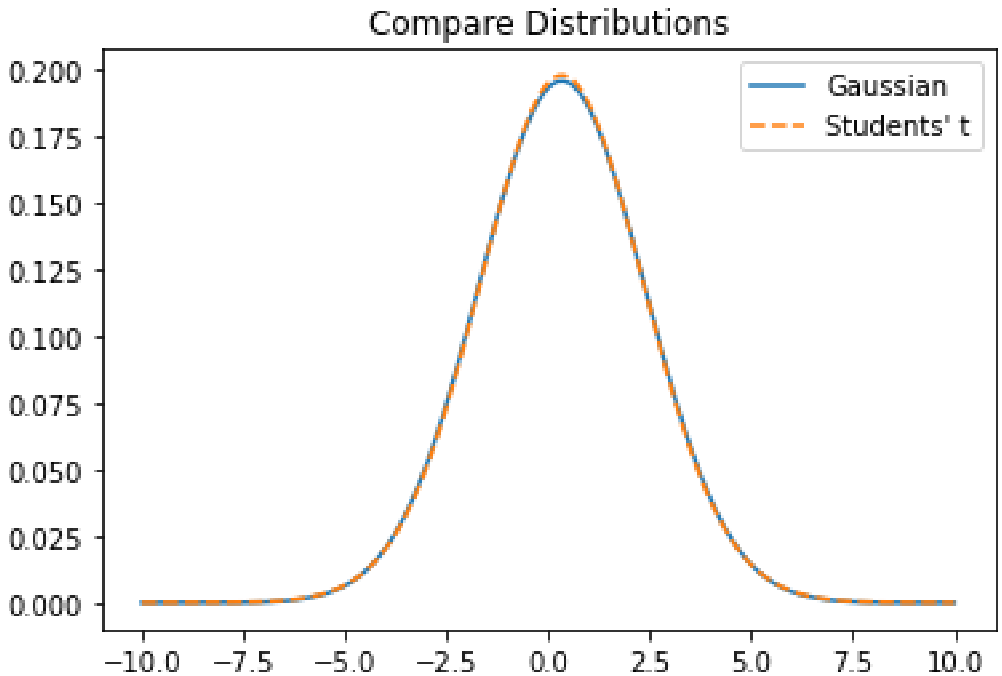


3. [10 points] Randomly draw 30 samples from a Gaussian distribution $\mathcal{N}(0, 2)$. Use the 30 samples as your observations to find the maximum likelihood estimation (MLE) for a Gaussian distribution and a student t distribution. For both distributions, please use L-BFGS to optimize the parameters. For student t, you need to estimate the degree of the freedom as well. Draw a plot of the estimated the

Gaussian distribution density, student t density and the scatter data points. What do you observe, and why? Next, we inject three noises into the data: we append $\{8, 9, 10\}$ to the 30 samples. Find the MLE for the Gaussian and student t distribution again. Draw the density curves and scatter data points in another figure. What do you observe, and why?

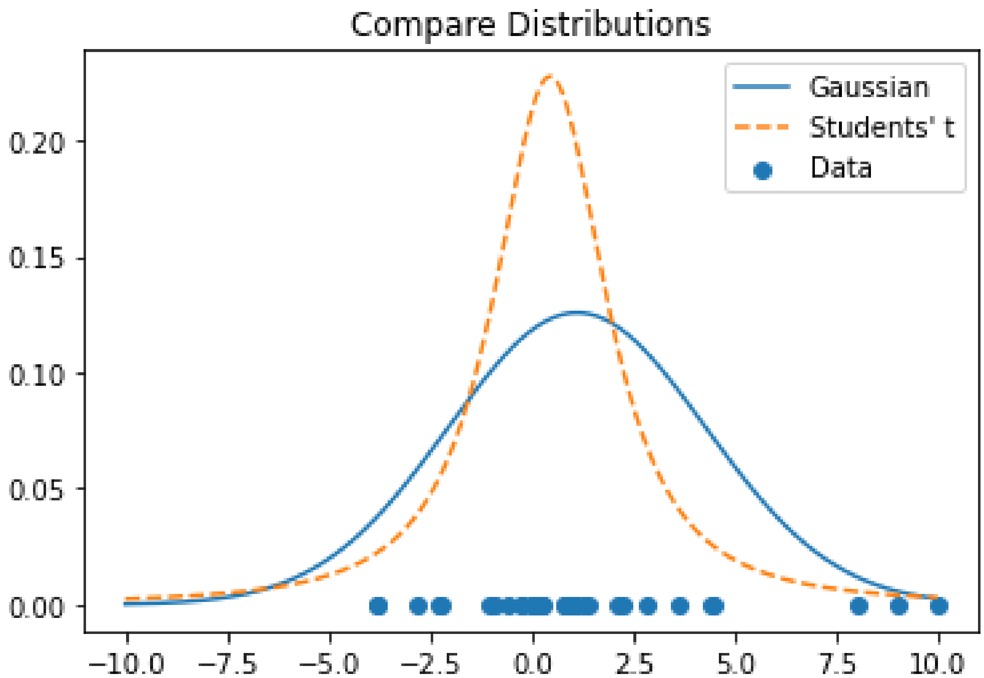
Answer

From initially sampled points, we get very identical plots for both of the distributions (shown below).



However when additional points are added, Gaussian distribution becomes a lot flatter however students t distribution maintains it's shape. This is important because the behaviour of this distribution is not affected by outliers and should be good if there's noise in data. In the plot below, I've also plotted the data points along x axis where we can see how they're distributed.

Code can be found **here**.


Compare Distributions
Gaussian
Students' t
Data
0.20
0.15
0.10
0.05
0.00
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0