Homework 3 Solutions

1 Probability and generative modeling

1. *Smartphone factories.*

   (a) Let A_i denote the event that the smartphone came from factory i for $i = 1$, 2, or 3. Let B denote the event that the smartphone is defective. Then our goal is to compute $\Pr(B)$.

   $$\Pr(B) = \sum_{i=1,2,3} \Pr(A_i \cap B) = \sum_{i=1,2,3} \Pr(A_i)\Pr(B \mid A_i) = 0.5 \cdot 0.1 + 0.25 \cdot 0.2 + 0.25 \cdot 0.24 = \frac{4}{25} = 0.16$$

   (b) Using the notation from part (a), our goal is to compute $\Pr(A_i \mid B)$ for $i = 1$, 2, or 3. Applying Bayes' rule, we have
   $$\Pr(A_i \mid B) = \frac{\Pr(A_i)\Pr(B \mid A_i)}{\Pr(B)}.$$
   Applying our result from part (a), we have
   $$\Pr(A_1 \mid B) = \frac{0.5 \cdot 0.1}{0.16} = \frac{5}{16} = 0.3125$$
   $$\Pr(A_2 \mid B) = \frac{0.25 \cdot 0.2}{0.16} = \frac{5}{16} = 0.3125$$
   $$\Pr(A_3 \mid B) = \frac{0.25 \cdot 0.24}{0.16} = \frac{6}{16} = 0.375$$

2. *The doctor's office.* Let A denote the event that he has the flu and B denote the event that he has a fever. Our goal is to compute $\Pr(A \mid B)$. We can first work out $\Pr(B)$.

   $$\Pr(B) = \Pr(A \cap B) + \Pr(\neg A \cap B) = \Pr(A)\Pr(B \mid A) + \Pr(\neg A)\Pr(B \mid \neg A) = 0.25 \cdot 0.75 + 0.75 \cdot 0.5 = \frac{9}{16} = 0.5625.$$

   Now applying Bayes' rule gives us

   $$\Pr(A \mid B) = \frac{\Pr(A)\Pr(B \mid A)}{\Pr(B)} = \frac{0.25 \cdot 0.75}{0.5625} = \frac{1}{3}$$

3. *Minimum of random variables*

   (a) Consider the following table in which the (i, j) entry corresponds to the value of X when $X_1 = i$ and $X_2 = j$.

	$X_2 = 1$	$X_2 = 2$	$X_2 = 3$	$X_2 = 4$	$X_2 = 5$	$X_2 = 6$
$X_1 = 1$	1	1	1	1	1	1
$X_1 = 2$	1	2	2	2	2	2
$X_1 = 3$	1	2	3	3	3	3
$X_1 = 4$	1	2	3	4	4	4
$X_1 = 5$	1	2	3	4	5	5
$X_1 = 6$	1	2	3	4	5	6

   Since each of these entries occurs with probability 1/36, we only need to count the number of times a particular number, say 3, occurs and divide by 36 to get the probability that $X = 3$. Thus we have the following.

$\Pr(X=1)$	$\Pr(X=2)$	$\Pr(X=3)$	$\Pr(X=4)$	$\Pr(X=5)$	$\Pr(X=6)$
$11/36$	$9/36 = 1/4$	$7/36$	$5/36$	$3/36 = 1/12$	$1/36$

(b) We can use the formula for expectation to get:

$$\mathbb{E}[X] = \sum_{x=1}^{6} x \Pr(X=x) = 1 \cdot \frac{11}{36} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{7}{36} + 4 \cdot \frac{5}{36} + 5 \cdot \frac{1}{12} + 6 \cdot \frac{1}{36} = \frac{91}{36}$$

(c) To compute the variance, we first compute

$$\mathbb{E}[X^2] = \sum_{x=1}^{6} x^2 \Pr(X=x) = 1 \cdot \frac{11}{36} + 2^2 \cdot \frac{1}{4} + 3^2 \cdot \frac{7}{36} + 4^2 \cdot \frac{5}{36} + 5^2 \cdot \frac{1}{12} + 6^2 \cdot \frac{1}{36} = \frac{301}{36}$$

Then $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{301}{36} - \left(\frac{91}{36}\right)^2 = \frac{2555}{1296}$ and $\mathrm{std}(X) = \sqrt{\mathrm{Var}(X)} = \sqrt{\frac{2555}{1296}}$.

4. *Independence*

   (a) Independent. The rank of a card has no bearing on its suit.

   (b) Not independent. The probability that a random card is a spade is $1/4$. But if our first draw is a spade, the probability that the second card is also a spade drops to $12/51$.

5. *Correlation*

   (a) Positively correlated. Rainstorms often last several days.

   (b) Negatively correlated. Generally, beachgoing weather is not conducive for skiing, and vice versa.

   (c) Uncorrelated. Social security numbers are not assigned sequentially.

6. *Gaussian parameters*

   (a) Based on the info given, we already know that $\mu = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$. Thus we just need to compute $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$.

$$\begin{aligned} \Sigma_{11} &= \mathrm{var}(x) = \mathrm{std}(x)^2 = 1 \\ \Sigma_{22} &= \mathrm{var}(y) = \mathrm{std}(y)^2 = \frac{1}{4} \\ \Sigma_{12} = \Sigma_{21} &= \mathrm{cov}(x,y) = \mathrm{corr}(x,y) \cdot \mathrm{std}(x) \cdot \mathrm{std}(y) = -\frac{1}{2} \cdot 1 \cdot \frac{1}{2} = -\frac{1}{4} \end{aligned}$$

   Thus $\Sigma = \begin{pmatrix} 1 & -0.25 \\ -0.25 & 0.25 \end{pmatrix}$.

   (b) We again can easily see $\mu = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. We can also write

$$\begin{aligned} \Sigma_{11} &= \mathrm{var}(x) = \mathrm{std}(x)^2 = 1 \\ \Sigma_{22} &= \mathrm{var}(y) = \mathrm{var}(x) = 1 \\ \Sigma_{12} = \Sigma_{21} &= \mathrm{cov}(x,y) = \mathrm{cov}(x,x) = \mathrm{var}(x) = 1 \end{aligned}$$

   Thus $\Sigma = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

7. *Gaussian contour lines* The contour lines should look something like this.

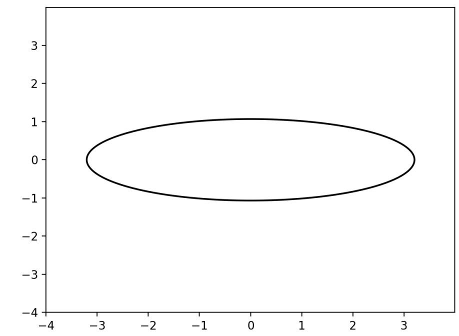

Figure 1: Contour line for (a)

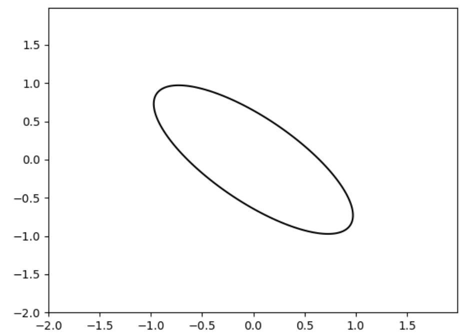

Figure 2: Contour line for (b)

8. *Gaussian contour lines* The sample scatter plots should look something like this.

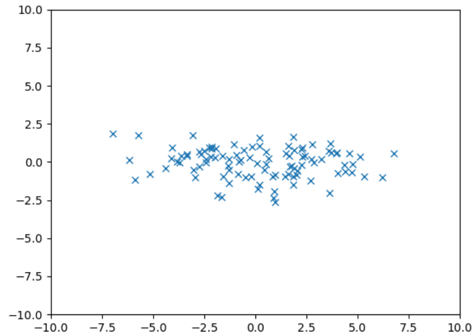

Figure 3: Scatter plot for (a)

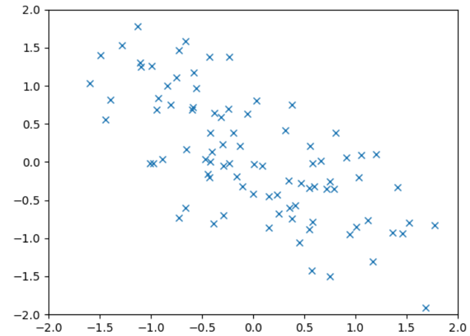

Figure 4: Scatter plot for (b)

9. *Optimal classifiers* Based on the densities of P_1, P_2, P_3, we can see that the optimal classifier will depend on whether or not a given point x is positive or negative. In particular, we need to compute $\pi_j P_j(x)$ for each classifier $j = 1, 2, 3$ for the case when x is positive and the case when x is negative. The following table shows these calculations

	$x \leq 0$	$x > 0$
$\pi_1 P_1(x)$	$\frac{1}{3} \cdot \frac{7}{8} = \frac{7}{24}$	$\frac{1}{3} \cdot \frac{1}{8} = \frac{1}{24}$
$\pi_2 P_2(x)$	$\frac{1}{6} \cdot 0 = 0$	$\frac{1}{6} \cdot 1 = \frac{1}{6}$
$\pi_3 P_3(x)$	$\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$	$\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

Based on these values, we have

$$h^*(x) = \begin{cases} \text{Class 1} & \text{if } x \leq 0 \\ \text{Class 3} & \text{if } x > 0 \end{cases}$$

2 Linear algebra

1. $(1/\sqrt{14}, 2/\sqrt{14}, 3/\sqrt{14})$
2. $(-1/\sqrt{2}, 1/\sqrt{2})$ and $(1/\sqrt{2}, -1/\sqrt{2})$
3. $x \cdot x = 25 \Leftrightarrow \|x\| = 5$. All points of length 5: a sphere, centered at the origin, of radius 5.
4. $f(x) = 2x_1 - x_2 + 6x_3 = w \cdot x$ for $w = (2, -1, 6)$.

5. A is 10×30 and B is 30×20

6. (a) X is $n \times d$
   (b) XX^T is $n \times n$
   (c) $(XX^T)_{ij} = x^{(i)} \cdot x^{(j)}$

7. $((x^T x)(x^T x)(x^T x)) = (\|x\|^2)^3 = 10^6$

8. $x^T x = \|x\|^2 = 35$ and

$$x^T x = \begin{bmatrix} 1 & 3 & 5 \\ 3 & 9 & 15 \\ 5 & 15 & 25 \end{bmatrix}$$

9. The angle θ between x and y satisfies $\cos\theta = x^T y / \|x\|\|y\| = 1/2$, so θ is 60 degrees.

10.

$$M = \begin{bmatrix} 3 & 1 & -2 \\ 1 & 0 & 0 \\ -2 & 0 & 6 \end{bmatrix}$$

11. *Symmetric Matrices*
   (a) $(AA^T)^T = (A^T)^T A^T = AA^T$, Thus AA^T is symmetric.
   (b) $(A^T A)^T = A^T (A^T)^T = A^T A$, Thus $A^T A$ is symmetric.
   (c) $(A + A^T)^T = (A^T + A) = (A + A^T)$, Thus $(A + A^T)$ is symmetric
   (d) $(A - A^T)^T = (A^T - A) \neq (A - A^T)$, Thus $(A - A^T)$ need not be symmetric

12. (a) $|A| = 8! = 40320$
   (b) $A^{-1} = \text{diag}(1, 1/2, 1/3, 1/4, 1/5, 1/6, 1/7, 1/8)$

13. *Orthonormal matrices*
   (a) UU^T is the identity matrix
   (b) $U^{-1} = U^T$

14. Since A is singular matrix, $|A| = 0 \implies z - 6 = 0 \implies z = 6$