Feng Wei

wei8

ECE 8560 Takehome#1

Bayesian Classifier Design and Implementation

1. Engineering decisions and associated rational

We have know that this is a $C = 3$ and $d = 4$ problem. For the training data we know that the first 5000 samples correspond to $Class_1$, the second 5000 samples correspond to $Class_2$, and the third 5000 samples correspond to $Class_3$. That's a good news. But we don't know the exact distribution function for each class, that is a bad news.

Since the samples are $d = 4$ vectors, we can't plot them in a figure to see the shape of their distribution and estimate their probability density function.

Though we can't plot a $d = 4$ vector in a figure, we can plot each dimension of the vector in different figures. We could estimate the distribution of each dimension first, and then estimate the distribution of the samples.

- For $Class_1$

We know that samples are vectors like $\bar{x}_i = (v_0, v_1, v_2, v_3)$. So we first plot the histogram of v_0.

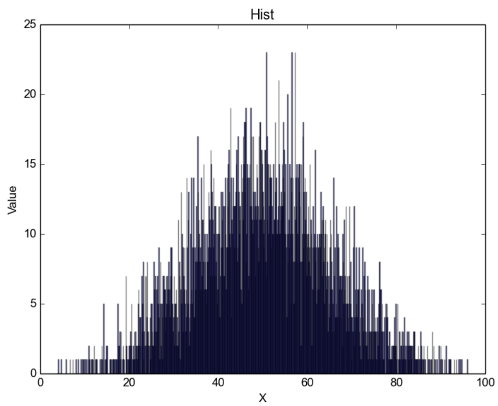

Figure 1(Histogram of v_0 for $Class_1$)

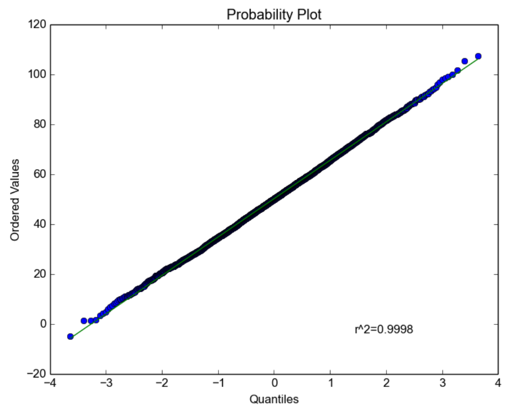

Figure 2(Q-Q plot)

From figure 1 we can see the distribution of v_0 like a normal distribution, now we want to know the relationship between normal distribution and the distribution of v_0.

In statistics, a Q–Q plot("Q" stands for quantile) is a probability plot, which is a graphical method for comparing two probability distributions by plotting their quantiles against each other.(From wiki) So we plot a Q-Q plot for v_0 and normal distribution.

From figure 2 we can see v_0 and normal distribution nearly the same, so we approximate $v_0 \sim N(\mu, \sigma)$.

For v_1, v_2, v_3 we do the same step as v_0 (See Appendix). And we find out that they all nearly have the same distribution as normal distribution. That means $v_i \sim N(\mu_i, \sigma_i)$.

Now we approximate that each dimension of the samples in $Class_1$ is a normal distribution. So we want to know if the samples' distribution a Multivariate normal distribution (Multivariate Gaussian distribution)

We compute the $\bar{\mu}$ and Σ of $Class_1$:

$$Cov_1 = \begin{bmatrix} 236.195941 & 3.92567954 & 6.67486635 & 14.8315462 \\ 3.92567954 & 218.9982569 & 6.37804829 & -5.58231856 \\ 6.67486635 & 6.37804829 & 224.4166463 & 1.46498982 \\ 14.83154628 & -5.58231856 & 1.46498982 & 648.6540535 \end{bmatrix}$$

$$Mean_1 = \begin{bmatrix} 50.11712203 & -4.97038793 & -24.81182102 & -49.81198585 \end{bmatrix}$$

In Cov_1 we see that $\Sigma_{i,j} \sim 0$ if $i \neq j$, so we approximate Cov_1 to a diagonal matrix like below:

$$Diag\_Cov_1 = \begin{bmatrix} \sigma_{11}^2 & 0 & 0 & 0 \\ 0 & \sigma_{22}^2 & 0 & 0 \\ 0 & 0 & \sigma_{33}^2 & 0 \\ 0 & 0 & 0 & \sigma_{44}^2 \end{bmatrix}$$

Now we could approximate that Samples in $Class_1$ is a Multivariate Gaussian Distribution with $\bar{\mu} = Mean_1$ and $\Sigma = Diag\_Cov_1$.

- For $Class_2$

  compute the $\bar{\mu}$ and Σ of $Class_2$:

$$Cov_2 = \begin{bmatrix} 2454.81251 & -19.4018922 & -1.03424872 & -8.11083884 \\ -19.4018922 & 626.827211 & 2.12538601 & 10.5259739 \\ -1.03424872 & 2.12538601 & 25.5821521 & 3.23909365 \\ -8.11083884 & 10.5259739 & 3.23909365 & 1610.92477 \end{bmatrix}$$

$$Mean_2 = \begin{bmatrix} 24.13111767 & -0.11837553 & -25.04828746 & 0.29147143 \end{bmatrix}$$

Do the same step as $Class_1$. And we approximate that Samples in $Class_2$ is a Multivariate Gaussian Distribution with $\bar{\mu} = Mean_2$ and $\Sigma = Diag\_Cov_2$.

- For $Class_3$

  compute the $\bar{\mu}$ and Σ of $Class_3$:

$$Cov_3 = \begin{bmatrix} 642.931240 & -28.3620882 & 3.27287974 & 2.99507042 \\ -28.3620882 & 2486.55244 & -20.9724965 & -1.76070235 \\ 3.27287974 & -20.9724965 & 628.685514 & -1.88547571 \\ 2.99507042 & -1.76070235 & -1.88547571 & 24.5768384 \end{bmatrix}$$

$$Mean_3 = \begin{bmatrix} 49.70218541 & 5.40154144 & 24.55009373 & -49.93734216 \end{bmatrix}$$

Do the same step as $Class_1$. And we approximate that Samples in $Class_3$ is a Multivariate Gaussian Distribution with $\bar{\mu} = Mean_3$ and $\Sigma = Diag\_Cov_3$

2. Show the exact form of the discriminant function used for each class.

- Gaussian Model

$$p(\bar{x}) = (2\pi)^{-\frac{d}{2}} \left|\Sigma_i\right|^{-\frac{1}{2}} exp[-\frac{1}{2}(\bar{x}-\bar{\mu_i})^T \Sigma_i^{-1}(\bar{x}-\bar{\mu_i})]$$

- Bayes Rules

$$P(\omega_i|\bar{x}) = p(\bar{x}|\omega_i) * P(\omega_i)/p(\bar{x}) \qquad p(\bar{x}) = \Sigma p(\bar{x}|\omega_i)$$

- Discriminant Functions

Define a discriminant function for the ith class as below:

$$g_i(\bar{x}) = P(w_i|\bar{x})$$

In that case we know that each class have the same probability. That means:

$$P(\omega_1) = P(\omega_2) = P(\omega_3) = \frac{1}{3}$$

Now we want to find the largest discriminant function, we have know that they have the same priori probability. So choose the class for which $p(\bar{x}|\omega_i)$ is largest. Hence the log function is a monotonically increasing of $g_i(\bar{x})$, we set a alternative discriminant function:

$$g_i^{'}(\bar{x}) = log\{p(\bar{x}|\omega_i)\}$$

$$g_i^{'}(\bar{x}) = -\frac{1}{2}(\bar{x}-\bar{\mu}_i)^T\Sigma_i^{-1}(\bar{x}-\bar{\mu_i}) - (\frac{d}{2})log(2\pi) - \frac{1}{2}|\Sigma_i|$$

For each class $\bar{\mu_i} = Mean_i$ and $\Sigma_i = Diag\_Cov_i$

$$Mean_1 = \begin{bmatrix}50.11712203 & -4.97038793 & -24.81182102 & -49.81198585\end{bmatrix}$$

$$Mean_2 = \begin{bmatrix}24.13111767 & -0.11837553 & -25.04828746 & 0.29147143\end{bmatrix}$$

$$Mean_3 = \begin{bmatrix}49.70218541 & 5.40154144 & 24.55009373 & -49.93734216\end{bmatrix}$$

$$Diag\_Cov_2 = \begin{bmatrix}2454.81250858 & 0 & 0 & 0\\ 0 & 626.82721073 & 0 & 0\\ 0 & 0 & 25.58215206 & 0\\ 0 & 0 & 0 & 1610.9247698\end{bmatrix} Diag\_Cov_1 = \begin{bmatrix}236.19594156 & 0 & 0 & 0\\ 0 & 218.99825698 & 0 & 0\\ 0 & 0 & 224.41664637 & 0\\ 0 & 0 & 0 & 648.65405355\end{bmatrix}$$

$$Diag\_Cov_3 = \begin{bmatrix}642.9312398 & 0 & 0 & 0\\ 0 & 2486.55244034 & 0 & 0\\ 0 & 0 & 628.68551414 & 0\\ 0 & 0 & 0 & 24.57683837\end{bmatrix}$$

3. Estimate your $P(error)$, using the training data with known class.

Class\Assigned	C_1	C_2	C_3	$Total$
C_1	4542	314	144	5000
C_2	578	4389	33	5000
C_3	234	49	4717	5000

$$P(error) = P(\bar{x} \ is \ assigned \ to \ the \ wrong \ class)$$

$$P(error) = \Sigma\Sigma P(C_j|C_i) * P(C_i) \quad where \ i \neq j; \ i, \ j \ from \ 1 \ to \ 3$$

$$P(error) = \frac{314 + 144 + 578 + 33 + 234 + 49}{5000 + 5000 + 500} = 9.01\%$$

Final result

$$Class = \begin{cases} Class_1 & g_1(\bar{x}_i) = max(g_1(\bar{x}_i), g_2(\bar{x}_i), g_3(\bar{x}_i)) \\ Class_2 & g_2(\bar{x}_i) = max(g_1(\bar{x}_i), g_2(\bar{x}_i), g_3(\bar{x}_i)) \\ Class_3 & g_3(\bar{x}_i) = max(g_1(\bar{x}_i), g_2(\bar{x}_i), g_3(\bar{x}_i)) \end{cases}$$

$(Class_1 \ 5322, Class_2 \ 4749, Class_3 \ 4929)$