Q1)

Blue: $\{-4,0\}\ \{-3,1\}\ \{-2,-1\}\ \{-2,-2\}\ \{0,-4\}$

	-4	3.24	0	1.44
	-3	0.64	1	4.84
	-2	0.04	-1	0.04
	-2	0.04	-2	0.64
	0	4.84	-4	7.84
mu,sig2	-2.2	1.76	-1.2	2.96

$$\mu_B: \begin{bmatrix} -2.2 \\ -1.2 \end{bmatrix} \qquad \Sigma_B: \begin{bmatrix} 2.2 & -2.55 \\ -2.55 & 3.7 \end{bmatrix}$$

Red: $\{-2,1\}\ \{-1,0\}\ \{1,1\}\ \{3,2\}\ \{3,5\}\ \{4,3\}\ \{5,2\}$

	-2	14.8776	1	1.0000
	-1	8.1633	0	4.0000
	1	0.7347	1	1.0000
	3	1.3061	2	0.0000
	3	1.3061	5	9.0000
	4	4.5918	3	1.0000
	5	9.8776	2	0.0000
mu,sig2	1.857	5.83673	2	2.285714

$$\mu_A: \begin{bmatrix} 1.857 \\ 2.286 \end{bmatrix} \qquad \Sigma_A: \begin{bmatrix} 6.81 & 2.38 \\ 2.38 & 2.57 \end{bmatrix}$$

Normal: $p(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\}$

$$P(x|y=A) = \frac{1}{2\pi|\Sigma_A|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(x-\mu_A)^T\Sigma_A^{-1}(x-\mu_A)\right\} \qquad |\Sigma_A| = 11.84,\quad \Sigma_A^{-1} = \begin{bmatrix} 0.271 & -0.2 \\ -0.2 & 0.575 \end{bmatrix}$$

$$P(x|y=B) = \frac{1}{2\pi|\Sigma_B|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(x-\mu_B)^T\Sigma_B^{-1}(x-\mu_B)\right\} \qquad |\Sigma_B| = 1.637,\quad \Sigma_B^{-1} = \begin{bmatrix} 2.26 & 1.56 \\ 1.56 & 1.34 \end{bmatrix}$$

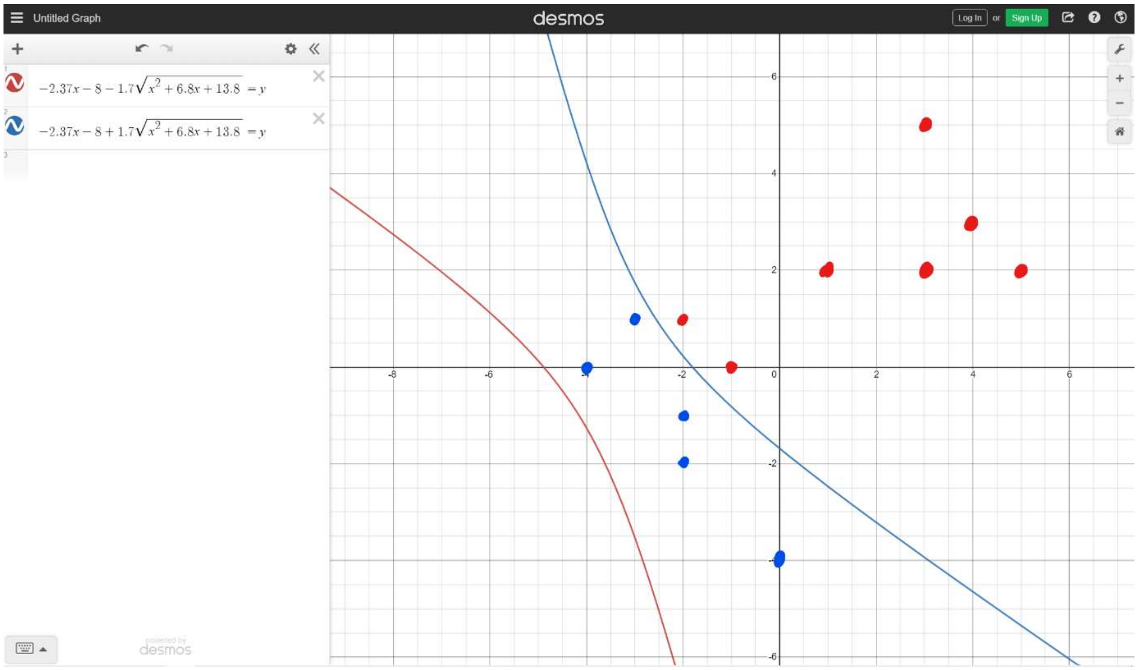
Bayes Rule: $\frac{P(x|y=A)}{P(x|y=B)} \gtrless \frac{P(y=B)}{P(y=A)} = \frac{5}{7}$

$$g(x) = \frac{\sqrt{1.637}}{\sqrt{11.84}} \exp\left\{-\frac{1}{2}[x-1.857 \quad y-2.286] \begin{bmatrix} 0.271 & -0.2 \\ -0.2 & 0.575 \end{bmatrix} \begin{bmatrix} x-1.857 \\ y-2.286 \end{bmatrix}\right\}$$

$$\rightarrow g(x) = 0.37\exp\{x^2 + 0.38y^2 + 1.8xy + 6.7x + 6.1y + 9.4\} \gtrless \frac{5}{7}$$

Hyperbolic equation:

$$-2.37x - 8 - 1.7\sqrt{x^2 + 6.8x + 13.8} < y < -2.37x - 8 + 1.7\sqrt{x^2 + 6.8x + 13.8}$$


Untitled Graph
desmos
Log In
or
Sign Up
$-2.37x-8-1.7\sqrt{x^2+6.8x+13.8}=y$
$-2.37x-8+1.7\sqrt{x^2+6.8x+13.8}=y$
-8
-6
-2
0
2
4
6
6
4
2
-2
-6
powered by
desmos

Q2)

Part a)

Decision boundary is at $\frac{a_1+a_2}{2}$.

$$P(error) = \int_{-\infty}^{\frac{a_1+a_2}{2}} p(\omega_2|x)dx + \int_{\frac{a_1+a_2}{2}}^{\infty} p(\omega_1|x)dx$$

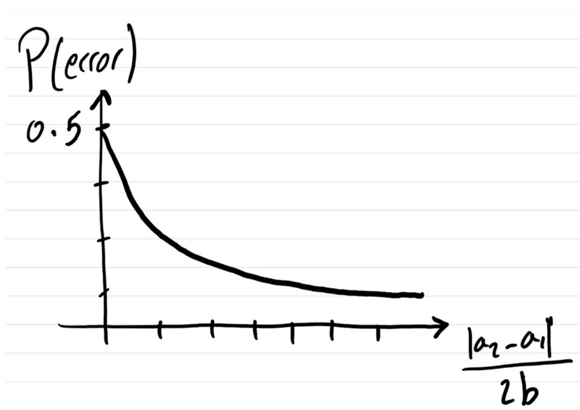
$$= \frac{1}{\pi b}\int_{-\infty}^{\frac{a_1+a_2}{2}} \frac{1/2}{1+\left(\frac{x-a_2}{b}\right)^2}dx + \frac{1}{\pi b}\int_{\frac{a_1+a_2}{2}}^{\infty} \frac{1/2}{1+\left(\frac{x-a_1}{b}\right)^2}dx$$

$$= \frac{1}{\pi b}\int_{-\infty}^{\frac{a_1+a_2}{2}} \frac{1}{1+\left(\frac{x-a_2}{b}\right)^2}dx = \frac{1}{\pi b}\int_{-\infty}^{\frac{a_1+a_2}{2}} \frac{1}{1+y^2}dy$$

$$= \frac{1}{\pi b}\left[\tan^{-1}\left|\frac{a_1-a_2}{2b}\right| - \tan^{-1}\infty\right] = \boldsymbol{\frac{1}{2} - \frac{1}{\pi}\tan^{-1}\left|\frac{a_1-a_2}{2b}\right|}$$

Part b)

The function is drawn bellow.



Part c)

It is obvious that $Max\ P(error) = 1/2$ and it happens at $\tan^{-1}\left|\frac{a_1-a_2}{2b}\right| = 0$ which occurs at $\left|\frac{a_1-a_2}{2b}\right| = 0$. This occurs at two situations:

1. When $\boldsymbol{a_1} = \boldsymbol{a_2}$ means the two distributions are the same.

2. When $\boldsymbol{b} \to \infty$ means both distributions are flat.

Q3)

Part a)

$$f(x_1,\dots,x_m|p_1,\dots,p_n) = \frac{n!}{\prod_1^m x_i!}\prod_1^n p_i^{x_i} \xrightarrow{iid} \log l(p) = \sum_1^m x_i\left(\log n! - \log\prod_1^m x_j!\right)$$

$$= \log n! - \sum_1^n \log x_i! + \sum_1^n x_i \log P_i$$

$$\sum_1^n P_i = 1 \to P_n = 1 - \sum_1^{n-1} P_i$$

$$\to \frac{P_i}{1-\sum_1^{n-1} P_i} = \frac{x_i}{x_n} \to \frac{P_i}{P_n} = \frac{x_i}{x_n} \to P_1 = \frac{x_1}{x_n}P_n \quad P_2 = \frac{x_2}{x_n}P_n \to P_i = \frac{x_i}{\sum_1^n x_i}$$

Part b)

Likelihood function $L(\omega) = P(X_1,\dots,X_d|\omega) = \prod_1^d P(X_i|\omega) = \prod_1^d \frac{1}{\omega} = \begin{cases}\left(\frac{1}{\omega}\right)^d & if\ \forall X_i \le \omega \\ 0 & Otherwise\end{cases}$

Part c)

MLE estimator $\arg\max_\omega L(\omega) = \arg\max_\omega P(X_1,\dots,X_d|\omega) = \arg\max_\omega \left(\frac{1}{\omega}\right)^d = \max\{X_1,\dots,X_d\}$

The optimal value of ω maximizes the MLE is X_{max}.

Part d)

$$P(x) = \frac{\lambda^x}{x!}\exp(-\lambda)$$

$$\to L(\lambda, x_1,\dots,x_n) = \prod_1^n \frac{\lambda^x}{x_d!}\exp(-\lambda)$$

$$\to \log L(\lambda) = -n\lambda - \sum_1^n \log(x_j!) + \log(\lambda)\sum_1^n x_j$$

$$\to \hat{\lambda} = \arg\max_\lambda P(\lambda, X_1,\dots,X_d)$$

$$\to \frac{d}{d\lambda} = -n + \frac{1}{\lambda}\sum_1^n x_j = 0$$

$$\to \hat{\lambda} = \frac{1}{n}\sum_1^n x_j$$

Therefore, the estimator λ is just the sample mean of the n observations in the sample.

Q4)

Part a)

$D = \{x_1, x_2, \ldots., x_d\}$ and $\mu \sim N(m_0, \Sigma_0)$

MAP estimation for μ:

$$L(\omega) = \prod_1^n p(x_k|\mu)$$

$$\to \log L(\mu) = \sum_1^n \log(p(x_k|\mu)) = -\frac{n}{2}\log\left((2\pi)^d|\Sigma|\right) - \sum_1^n \frac{1}{2}(x_k-\mu)^t\Sigma^{-1}(x_k-\mu)$$

$$\to p(\mu) = \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(\mu-m_0)^t\Sigma^{-1}(\mu-m_0)\right)$$

$\to$ MAP: $\hat{\mu} = \arg\max_\mu[\log L(\mu)] \times [p(\mu)]$

$$= \arg\max_\mu\left[-\frac{n}{2}\log\left((2\pi)^d|\Sigma|\right) - \sum_1^n\frac{1}{2}(x_k-\mu)^t\Sigma^{-1}(x_k-\mu)\right] \times \left[\frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(\mu - m_0)^t\Sigma^{-1}(\mu-m_0)\right)\right]$$

Part b)

$$\mu' = E\{X'\} = E\{AX\} = A\mu$$

$$\Sigma' = E\{(X'-\mu')(X'-\mu')^t\} = E\{(AX'-A\mu')(AX'-A\mu')^t\} = E\{A(X'-\mu')(X'-\mu')^tA^t\} = A\Sigma A$$

Likelihood $\to \log p(D'|X') = \log\prod_1^n p(x_k'|\mu') = \log\prod_1^n p(Ax_k|A\mu) = \sum_1^n\log\left(p(Ax_k|A\mu)\right)$

$$= -\frac{n}{2}\log[(2\pi)^d|A\Sigma A^t|] - \sum_1^n\frac{1}{2}(Ax_k-A\mu)^t(A\Sigma A^t)^{-1}(Ax_k-A\mu)$$

$$= -\frac{n}{2}\log[(2\pi)^d|A\Sigma A^t|] - \sum_1^n\frac{1}{2}((x-\mu)^tA^t)(A^{-1}\Sigma^{-1}A^{-1})\left(A(x_k-\mu)\right)$$

$$= -\frac{n}{2}\log[(2\pi)^d|A\Sigma A^t|] - \sum_1^n\frac{1}{2}(x_k-\mu)^t\Sigma^{-1}(x_k-\mu)$$

Density of μ' $\to p(\mu') = \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma_0'|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(\mu-m_0)^t\Sigma_0^{-1}(\mu-m_0)\right)$

$$\to \widehat{\mu'} = \arg\max_\mu\left[-\frac{n}{2}\log\left((2\pi)^d|\mathrm{A\Sigma A^t}|\right) - \sum_1^n\frac{1}{2}(x_k-\mu)^t\Sigma^{-1}(x_k-\mu)\right] \times \left[\frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma_0'|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(\mu-m_0)^t\Sigma_0^{-1}(\mu-m_0)\right)\right]$$

Two equations are the same. So, MAP estimate of μ is appropriate.