Problem 1

$P(G=1|X) = N(G=1;\mu_1,\Sigma) \equiv \frac{1}{(2\pi)^{\frac{p}{2}}|\Sigma|^{\frac{1}{2}}}e^{-\frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1)}$

$P(G=2|X) = N(G=2;\mu_2,\Sigma) \equiv \frac{1}{(2\pi)^{\frac{p}{2}}|\Sigma|^{\frac{1}{2}}}e^{-\frac{1}{2}(x-\mu_2)^T\Sigma^{-1}(x-\mu_2)}$

$log(P(G=1|X)) = -\frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1) - \frac{p}{2}log(2\pi) - \frac{1}{2}log(|\Sigma|)$

$log(P(G=2|X)) = -\frac{1}{2}(x-\mu_2)^T\Sigma^{-1}(x-\mu_2) - \frac{p}{2}log(2\pi) - \frac{1}{2}log(|\Sigma|)$

The discriminant function in the case when they are a multdimentional Guassian is given by

$\delta_1(x) = -\frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1) - \frac{p}{2}log(2\pi) - \frac{1}{2}log(|\Sigma|) + log(\pi_1)$

$\delta_2(x) = -\frac{1}{2}(x-\mu_2)^T\Sigma^{-1}(x-\mu_2) - \frac{p}{2}log(2\pi) - \frac{1}{2}log(|\Sigma|) + log(\pi_2)$

Since LDA corresponds to the case of equal covariance matrices the decision boundaries but with equal covariances. For decision purposes we can drop the two terms $-\frac{p}{2}ln(2\pi) - \frac{1}{2}ln(|\Sigma|)$ and use a discriminant $\delta(x)$ given by

$\delta_1(x) = -\frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1) + log(\pi_1)$

$\delta_2(x) = -\frac{1}{2}(x-\mu_2)^T\Sigma^{-1}(x-\mu_2) + log(\pi_2)$

Expanding the quadratic in the above expression we get:

$\delta_1(x) = -\frac{1}{2}(x^T\Sigma^{-1}x - x^T\Sigma^{-1}\mu_1 - \mu_1^T\Sigma^{-1}x + \mu_1^T\Sigma^{-1}\mu_1) + log(\pi_1)$

$\delta_2(x) = -\frac{1}{2}(x^T\Sigma^{-1}x - x^T\Sigma^{-1}\mu_2 - \mu_2^T\Sigma^{-1}x + \mu_2^T\Sigma^{-1}\mu_2) + log(\pi_2)$

Since $x^T\Sigma^{-1}x$ is a common term with the same value in all discriminant functions we can drop it and just consider the discriminant given by

$\delta_1(x) = \frac{1}{2}x^T\Sigma^{-1}\mu_1 + \frac{1}{2}\mu_1^T\Sigma^{-1}x - \frac{1}{2}\mu_1^T\Sigma^{-1}\mu_1 + log(\pi_1)$

$\delta_2(x) = \frac{1}{2}x^T\Sigma^{-1}\mu_2 + \frac{1}{2}\mu_2^T\Sigma^{-1}x - \frac{1}{2}\mu_2^T\Sigma^{-1}\mu_2 + log(\pi_2)$

Since $x^T\Sigma^{-1}\mu$ is a scalar, its value is equal to the value of its transpose so

$x^T\Sigma^{-1}\mu = (x^T\Sigma^{-1}\mu)^T = \mu^T(\Sigma^{-1})x = \mu^T\Sigma^{-1}x$, since Σ^{-1} is symmetric.

Thus the two linear terms in the above combine and we are left with

$\delta_1(x) = x^T\Sigma^{-1}\mu_1 - \frac{1}{2}\mu_1^T\Sigma^{-1}\mu_1 + log(\pi_1)$

$\delta_2(x) = x^T\Sigma^{-1}\mu_2 - \frac{1}{2}\mu_2^T\Sigma^{-1}\mu_2 + log(\pi_2)$

Then we can estimate π_k from data using $\pi_k = \frac{N_k}{N}$ for $k, i = 1, 2$ and we pick class 2 as the classification outcome if $\delta_2(x) > \delta_1(x)$. Thus,

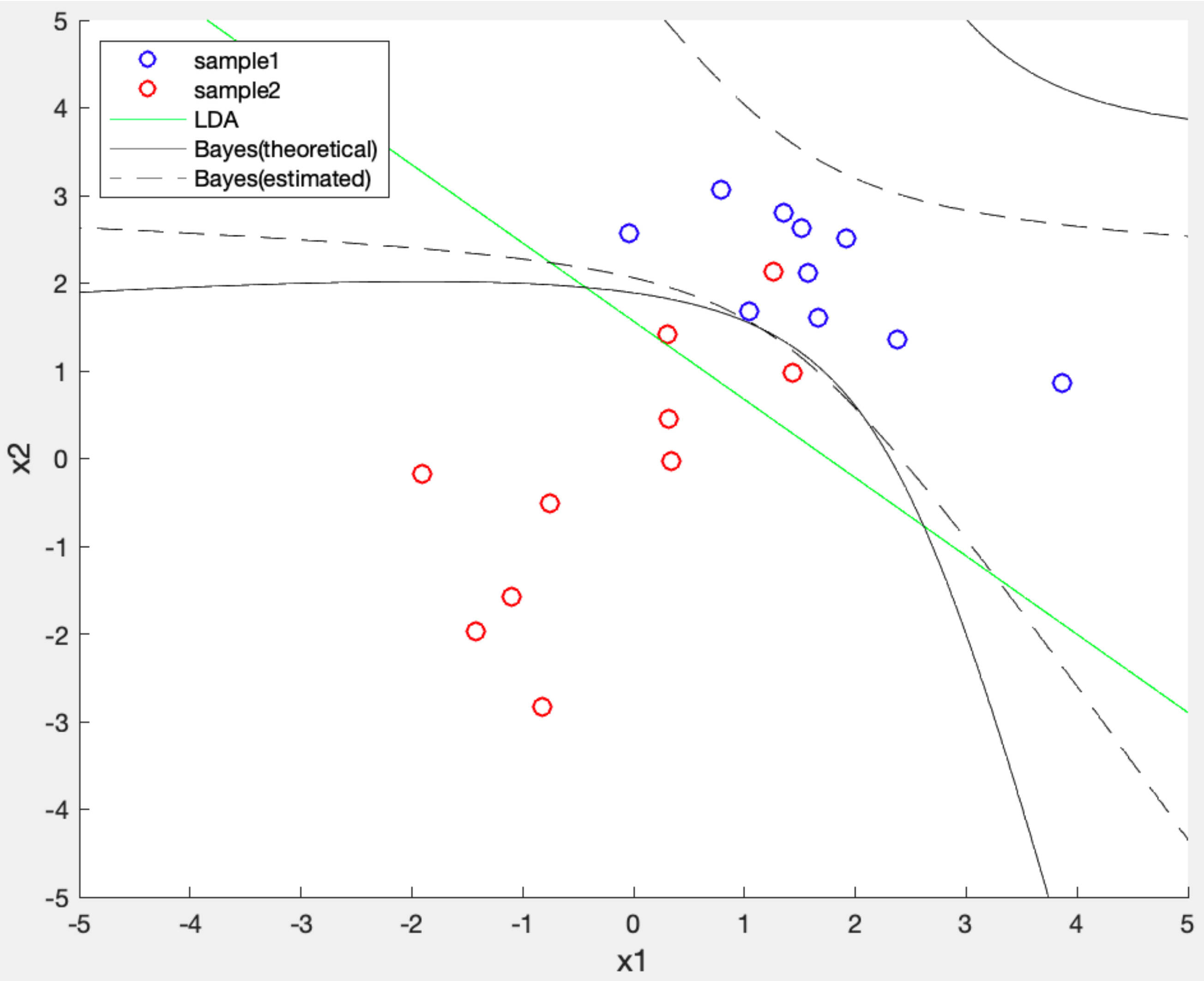
$x^T\Sigma^{-1}\mu_2 - \frac{1}{2}\mu_2^T\Sigma^{-1}\mu_2 + log(\frac{N_2}{N}) > x^T\Sigma^{-1}\mu_1 - \frac{1}{2}\mu_1^T\Sigma^{-1}\mu_1 + log(\frac{N_1}{N})$

which is

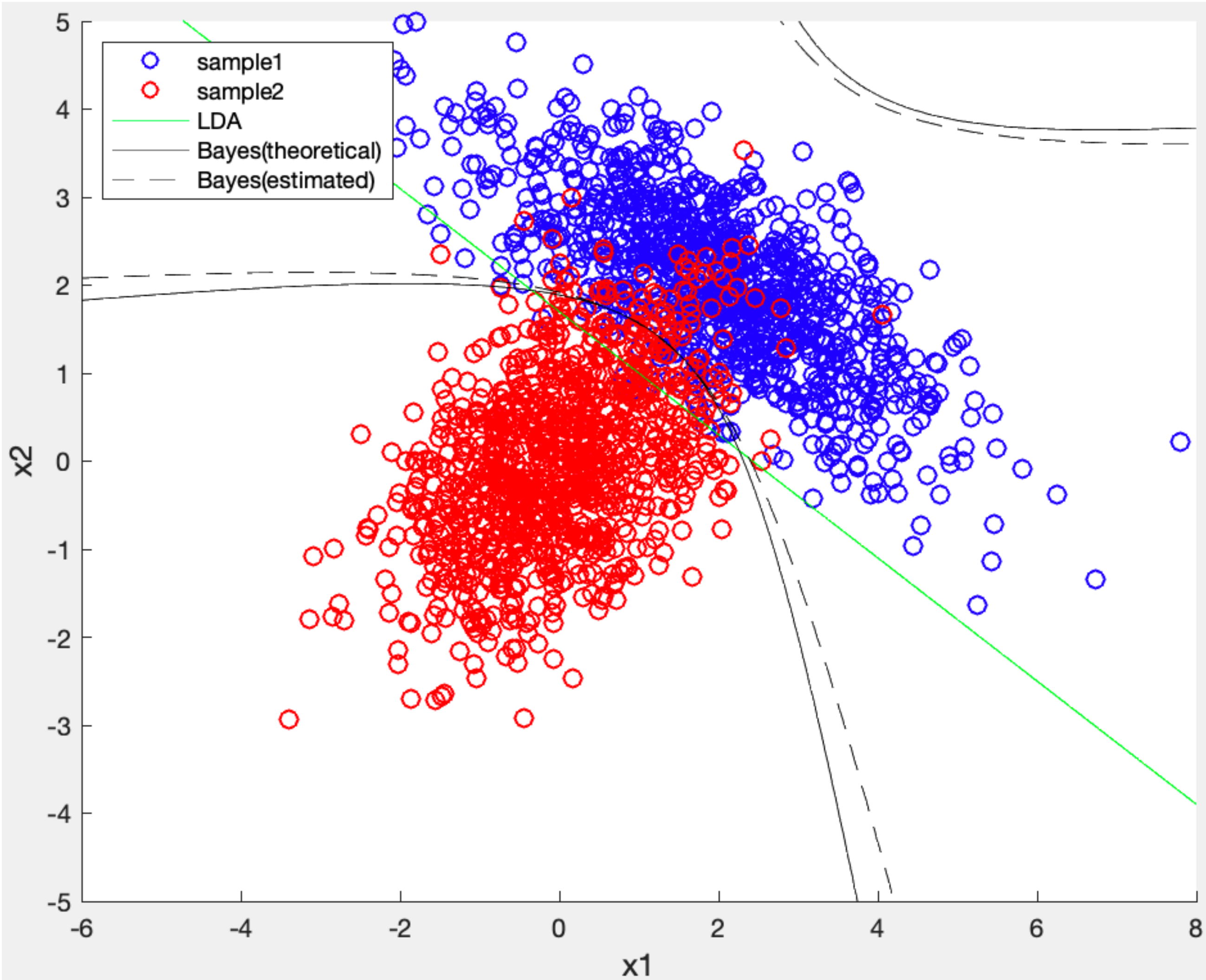
$x^T\Sigma^{-1}(\mu_2 - \mu_1) > \frac{1}{2}\mu_1^T\Sigma^{-1}\mu_2 - \frac{1}{2}\mu_1^T\Sigma^{-1}\mu_1 - log(\frac{N_2}{N_1})$

problem 2

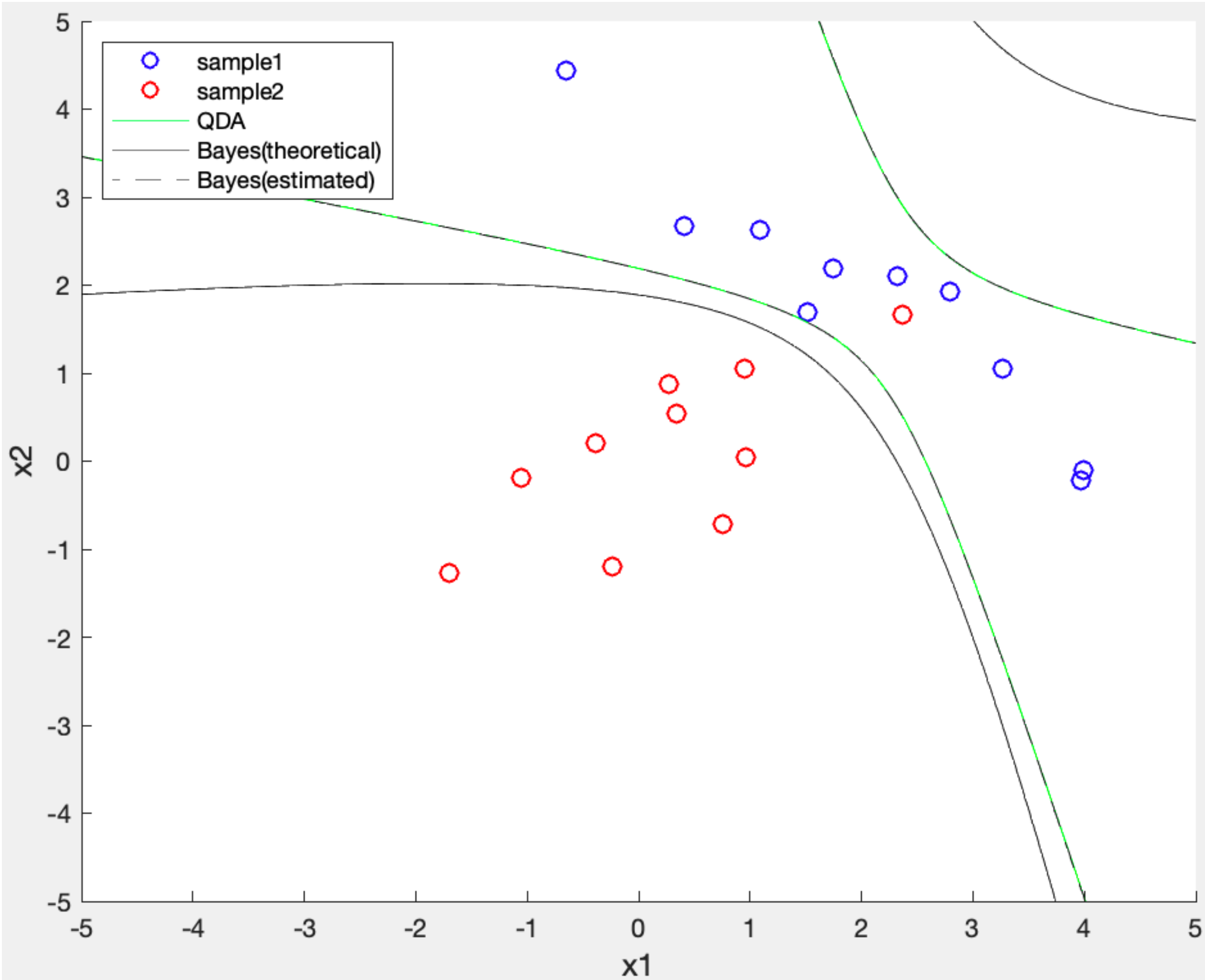
2.a.



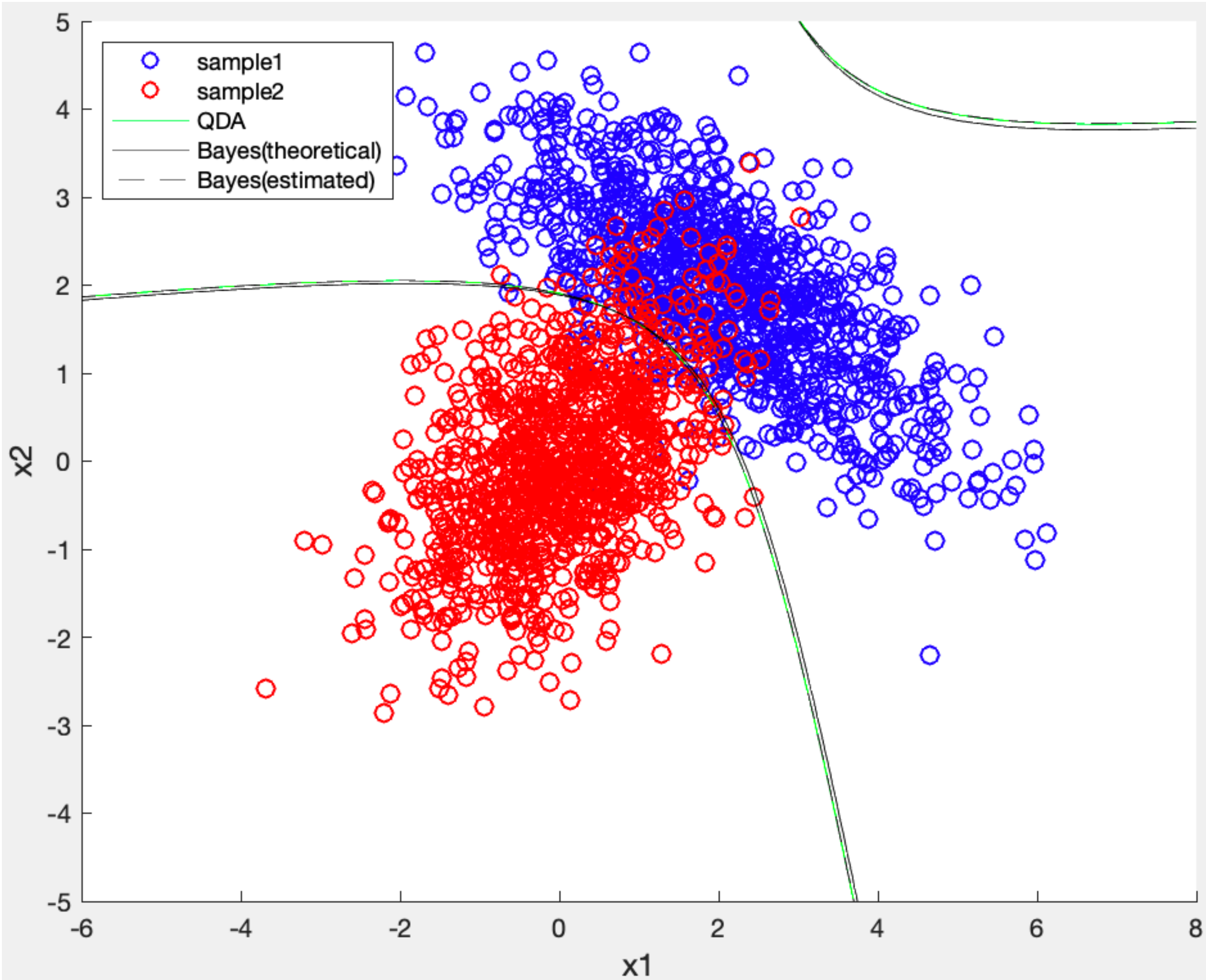
2.b.


sample1
sample2
LDA
Bayes(theoretical)
Bayes(estimated)
x1
x2
-6
-4
-2
0
2
4
6
8
5
4
3
2
1
0
-1
-2
-3
-4
-5

3.a.


sample1
sample2
QDA
Bayes(theoretical)
Bayes(estimated)
x1
x2
-5
-4
-3
-2
-1
0
1
2
3
4
5

3.b.



sample1
sample2
QDA
Bayes(theoretical)
Bayes(estimated)
x1
x2
-6
-4
-2
0
2
4
6
8
5
4
3
2
1
0
-1
-2
-3
-4
-5