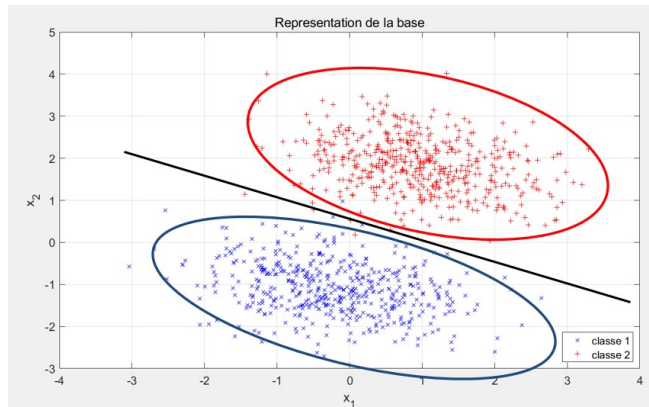


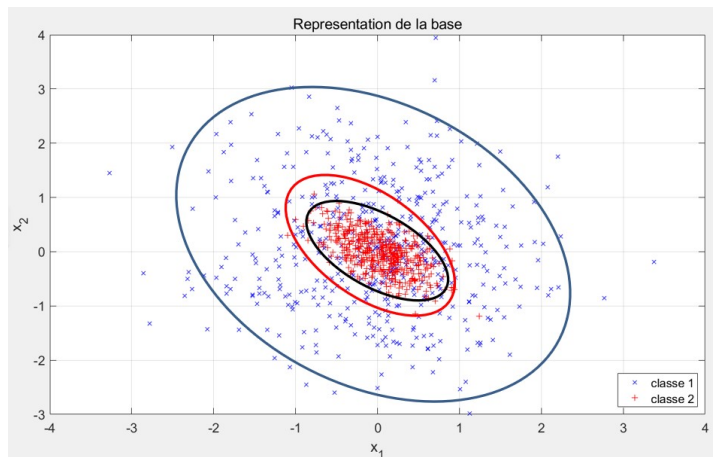
Audrey et Orlane

1.1



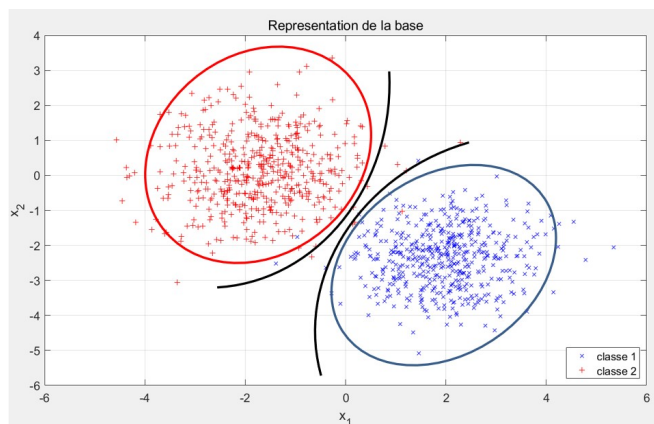
D'après la figure, l'allure des densités de probabilité sont des ellipses, les deux ellipses sont mieux distingués. On peut utiliser une ligne droite comme la frontière.

1.2



D'après la figure, l'allure des densités de probabilité sont des ellipses. Une ellipse est à l'intérieur d'une autre ellipse, on ne peut pas séparer complètement les deux. Donc on choisit une ellipse comme la frontière, bien qu'il existe un autre type de point dans ce classe, il est garanti que la probabilité d'erreur est minimal.

1.3



D'après la figure, l'allure des densités de probabilité sont des ellipses. Selon la position et la forme des deux ellipses, la ligne droite ne peut pas distinguer précisément, alors on choisit

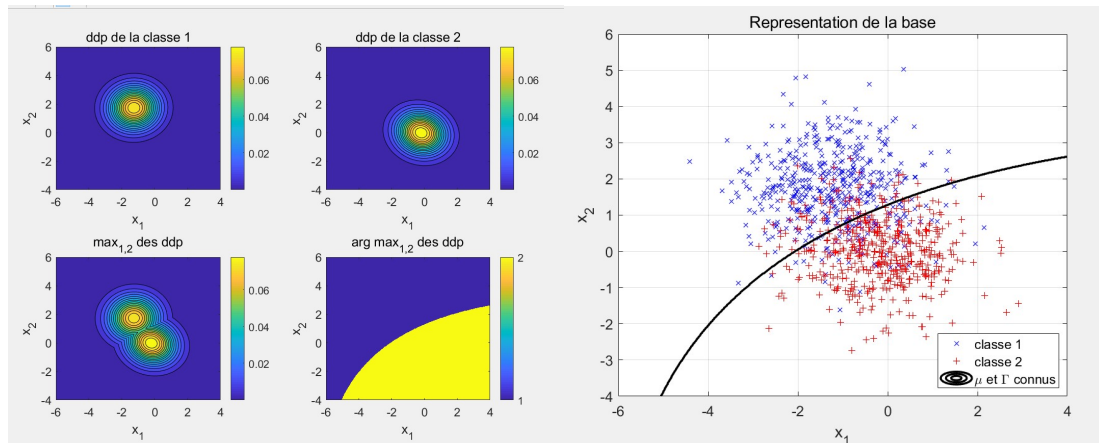
l'hyperbole comme la frontière.

1.4

D'après le cours, quand on connaît μ et Γ , on utilise le critère de la probabilité d'erreur minimale

$$P_{err} = P(\text{faire une erreur}) = \int_{\mathbb{R}^N} P(x, \text{faire une erreur}) dx$$

$$P_{err}^{opt} = \int \min(P(x, \omega_1), P(x, \omega_2)) dx$$



D'après la figure de ddp de la classe 1 et de la classe 2, on peut obtenir la figure de maximum des deux ddp. Et puis, d'après le formul (44), on obtient la figure de $\hat{d}_{\mu, \Gamma \text{ connus}}(x)$. C'est la surface discriminante.

$$\hat{d}_{\mu, \Gamma \text{ connus}}(x) = \arg \max_{c \in \{1, \dots, C\}} (f(x, (\mu_c, \Gamma_c))) \quad (44)$$

D'après le cours, l'équation suivie par la surface discriminante s'écrit comme :

$$\frac{1}{2} \ln \frac{|\Gamma_2|}{|\Gamma_1|} + \frac{1}{2} (Q_2(x) - Q_1(x)) \stackrel{\omega_1}{\omega_2} \geq 0$$

$$Q_c(x) = (x - \mu_c)^T \Gamma_c^{-1} (x - \mu_c)$$

Ici, d'après la figure, c'est presque une parabole.

En observant les graphiques de densité de probabilité dans différentes situations, on peut obtenir la forme approximative de la frontière. Ici, grâce au calcul précis, on peut obtenir la frontière qui minimise la probabilité d'erreur, et cette frontière est optimale. La frontière est une fonction quadratique, elle peut être décrite comme ellipse, hyperbole ou parabole. Dans la question précédente, on utilise aussi ces courbes pour exprimer la frontière. Cela est conforme à notre conclusion précédente.

2.1

D'après TP1, on a $\sigma = \sqrt{\frac{\mu_{\tau_g}(1 - \mu_{\tau_g})}{P_{gen}}}$. Quand on fait une seule fois l'expérience, on peut remplacer

$$\mu_{\tau_g} \text{ par } \tau_g. \text{ Donc, on a la barre d'erreur } \sigma = \sqrt{\frac{\tau_g(1 - \tau_g)}{P_{gen}}}.$$

Quand on exécute plusieurs fois le script, on a plusieurs de τ_g . On peut regarder que les

valeurs sont proches. Donc, on peut utiliser cette méthode pour obtenir la barre d'erreur.

```
tau_g (mu et Gamma connues) = 0.825
>> main
tau_g (mu et Gamma connues) = 0.88
>> main
tau_g (mu et Gamma connues) = 0.92
>> main
tau_g (mu et Gamma connues) = 0.85
```

2.2

```
>> main
tau_g (mu et Gamma connues) = 0.91 +- 0.0286
tau_g (lineaire) = 0.83 +- 0.0376
>>
```

Ici on regarde que la différence de τ_g entre les deux discriminateurs est 0.08, la barre d'erreur sont 0.0286 et 0.0376, qui sont beaucoup plus petit que 0.08. Donc dans ce cas là, on peut dire que le discriminateur de mu et Gamma connus est meilleur.

```
>> main
tau_g (mu et Gamma connues) = 0.86 +- 0.0347
tau_g (lineaire) = 0.83 +- 0.0376
```

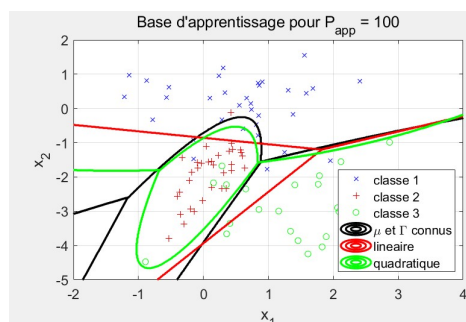
Et puis, on fait une nouvelle expérience, ici, la différence de τ_g entre les deux discriminateurs est 0.03, la barre d'erreur sont 0.0347 et 0.0376, qui sont plus grand que 0.03. Donc dans ce cas là, on ne peut pas dire que quel discriminateur est meilleur. Ici, on observe une perte de performance significative.

Donc, dans certain cas, on observe une perte de performance significative, il faut considérer cette phénomène.

2.3

```
>> main
tau_g (mu et Gamma connues) = 0.92 +- 0.0271
tau_g (lineaire) = 0.86 +- 0.0347
tau_g (quadratique) = 0.92 +- 0.0271
```

Ici, la différence de τ_g entre les deux discriminateurs est 0.06, la barre d'erreur sont 0.0347 et 0.0271, qui sont beaucoup plus petit que 0.06. Donc dans ce cas là, on peut dire que le discriminateur quadratique est meilleur. Et puis, la discriminateur de mu et Gamma connus est optimal, donc il est toujours le meilleur. Du coup, on observe des différences de performance significatives.



D'après la figure, la surface discriminant du discriminateur quadratique coïncide mieux celle du discriminateur de mu et Gamme connus. Donc son performance est meilleur. Mais pour le discriminateur linéaire, on suppose que les classes partagent la même matrice de covariance, donc la surface discriminant du discriminateur linéaire sont les lignes droites, du coup, il ne coïncide pas bien, son performance n'est pas bien.

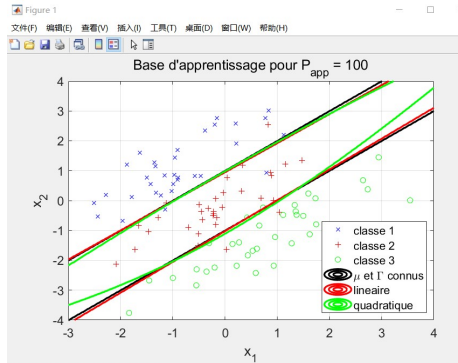
2.4

```

>> main
tau_g (mu et Gamma connus) = 0.9 +- 0.03
tau_g (lineaire) = 0.89 +- 0.0313
tau_g (quadratique) = 0.9 +- 0.03

```

Ici, la différence de τ_g entre les deux discriminateurs est 0.01, la barre d'erreur sont 0.0313 et 0.03, qui sont plus grand que 0.01. Donc dans ce cas là, on ne peut pas dire que quel discriminateur est meilleur. Ici, on observe une perte de performance significative.

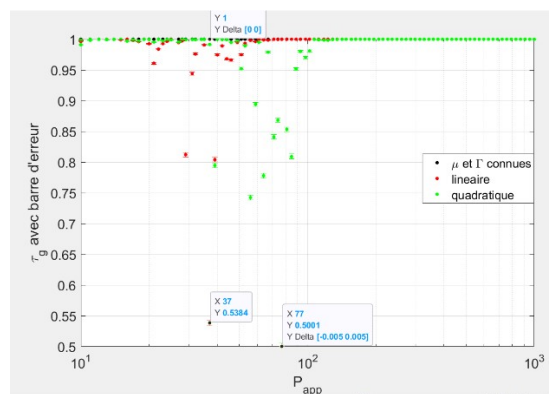


D'après la position des trois groups de point, ils peuvent être séparés par les lignes droites. Donc, la performance des deux discriminateurs sont près. Donc, on ne peut pas dire que quel discriminateur est meilleur. C'est pourquoi on observe une perte de performance significative.

2.5

C'est à dire que si dans un phénomène complexe, la performance du discriminateur quadratique est très bien, alors dans les phénomènes plus simples, il fonctionne aussi très bien. Pour le discriminateur linéaire, on suppose que les classes partagent la même matrice de covariance, mais pour le discriminateur quadratique, on ne suppose pas. Donc, le discriminateur quadratique peut gérer presque toutes les situations, qu'ils soient simples ou complexes, il dispose d'une bonne performance. Mais pour le discriminateur linéaire, dans certains cas particuliers, il peut avoir une bonne performance, mais dans les cas complexes, peut-être qu'il ne fonctionne pas très bien.

2.6



Pour bien observer la figure, il faut diminuer la barre d'erreur, donc on choisit un grand $P_{gen} = 10000$.

D'abord, le discriminateur de mu et Gamma connus est un discriminateur optimal, τ_g est toujours égale à 1.

Et puis, quand P_{app} est très grand ($P_{app} > 100$), les τ_g des trois discriminateurs sont

égales à 1. Les performances sont très bien.

D'après la figure, à côté de $P_{app} = 38$, il y a une réduction de τ_g du discriminateur linéaire. Ici, $N=38$, on fait une analogie comme TP1, quand on calcule les paramètres du discriminateur linéaire, si P_{app} est près que N , il y a une réduction de τ_g .

Pour le discriminateur quadratique, on regarde que à côté de $P_{app} = 76$, il y a aussi une réduction de τ_g . On pense que c'est parce que $P_{app} = 2N$.

Question : quel est le lien exact entre N , P_{app} et la réduction de τ_g ?