

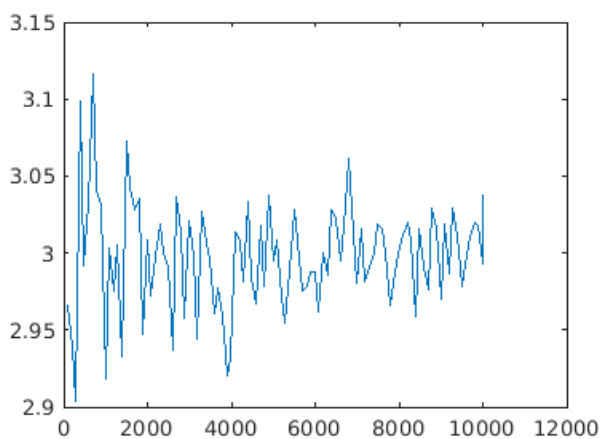
SI 221

Compte-rendu : TP décision bayésienne

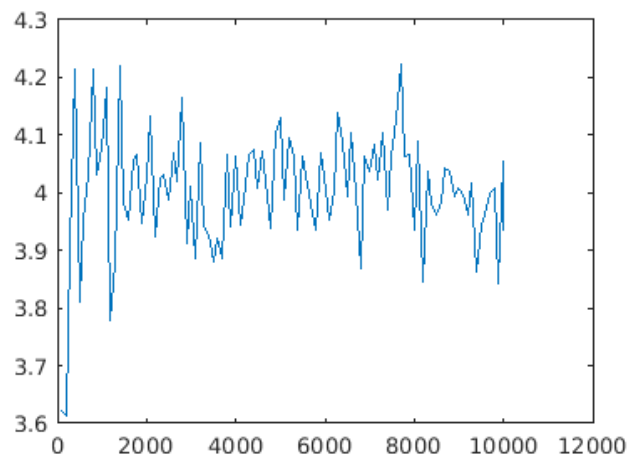
2. Génération d'une variable aléatoire gaussienne

Si X est une variable normale centrée, on sait que $Z = aX + b$ est une variable gaussienne de moyenne b et de variance a^2 . On utilise cette formule pour générer des échantillons de Z à partir de la fonction `randn` de Matlab.

En traçant la moyenne et la variance empirique en fonction du nombre d'échantillons N , on voit qu'on a besoin d'un nombre important d'échantillons pour avoir une précision acceptable, bien que ladite précision stagne à partir de $N = 3000$ environ. À gauche, on voit la moyenne en fonction de N , et à droite la variance.



Moyenne en fonction de N



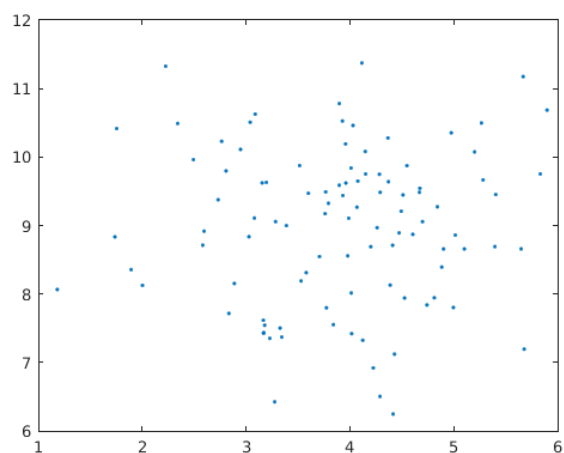
Variance en fonction de N

3. Génération de vecteurs aléatoires gaussiens

1) Similairement au cas unidimensionnel, le simple fait de rajouter un vecteur 2D m à la variable normale centrée crée une nouvelle variable gaussienne de moyenne m , donnant pour $m = [4 \ 9]$ le résultat ci-contre.

2) Si X est une variable normale centrée, l'expression $Z = SX + m$ donne un vecteur gaussien de moyenne m et de variance s , avec S étant la matrice diagonale dont les coefficients sont les racines carrées des coefficients de S . En effet :

$$\text{Var}(Z) = \text{Var}(SX) = S^2 \text{Var}(X) = s \text{Var}(X) \quad \text{car } S \text{ est diagonale.}$$
 Afficher les moyennes et variances empiriques de Z confirme ce résultat. Un exemple de tels échantillons est donné plus bas.



100 échantillons d'un vecteur 2D gaussien

3) On cherche à trouver U tel que :

$$\text{Var}(UX) = U^T U \text{Var}(X) = \Sigma \text{Var}(X)$$

Ainsi, il suffit de prendre $U = \sqrt{\Sigma}$ la racine carrée matricielle.

4) L'angle de l'ellipsoïde de Mahalanobis est donnée par :

$$s = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}$$

$$\alpha = \arctan\left(\frac{s_{12} s_{21}}{s_{11} - s_{22}}\right)$$

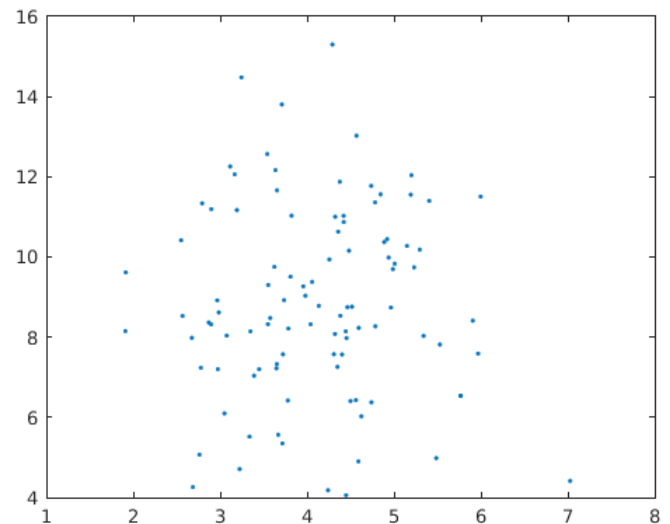
Ce qui nous donne $\alpha = \arctan(-4 / 3) = -0,4636$.

On vérifie à l'aide de la fonction eig que la matrice s vérifie bien la relation :

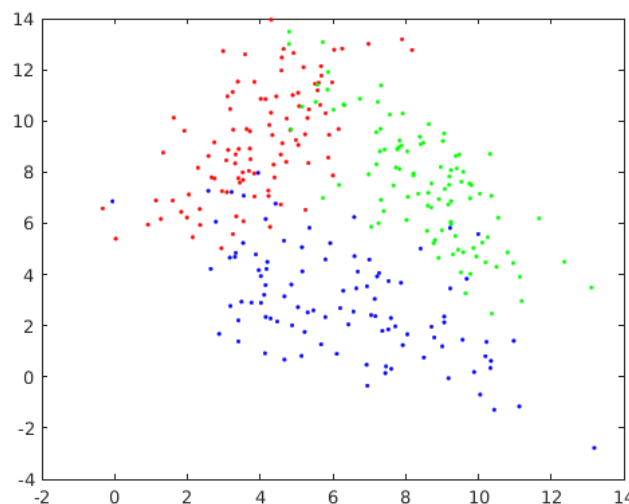
$$[V, Sd] = \text{eig}(s);$$

$$V * Sd * \text{transpose}(V) == s$$

5) L'affichage des échantillons des trois classes en les trois couleurs rouge, vert et bleu respectivement est donné ci-contre.



100 échantillons de $Z = SX + m$

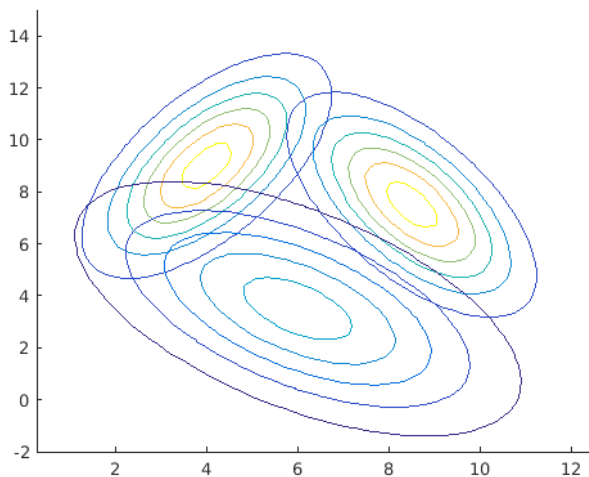


Trois classes de vecteurs gaussiens

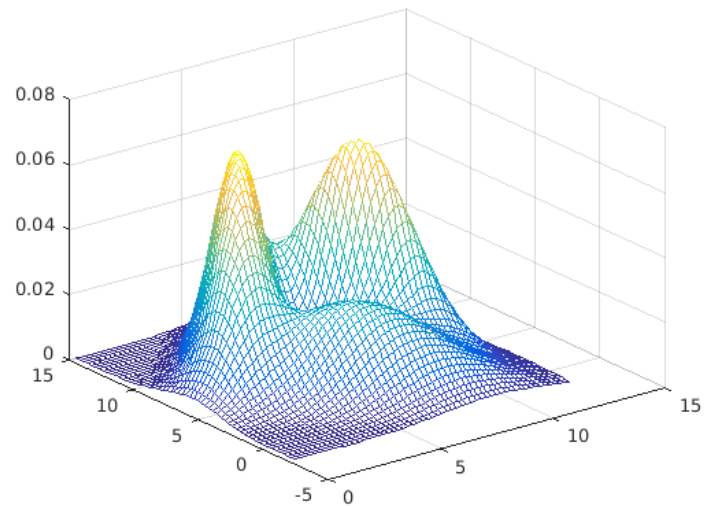
4. Courbes d'équidensité

1) 2) Pour chaque point de la grille, on calcule la densité de probabilité conditionnelle en ce point pour la classe 1. Cette expression est simplement $p_1(x|\omega_1) = \frac{1}{2\pi\sqrt{|s_1|}} e^{-\frac{(x-m_1)^T s_1^{-1} (x-m_1)}{2}}$, avec x parcourant chaque point de la grille.

3) 4) Les courbes d'équidensité prennent la forme d'ellipses centrées autour de la moyenne de la variable gaussienne liée à chaque classe. Dans le rendu 3D des trois fonctions de densité, on remarque que les 3 maxima locaux sont clairement visibles, disjoints, et ont une allure de cloche, ce qui est cohérent avec le fait que les fonctions de densité sont des gaussiennes. En effet, pour $x = m_i$, on voit que l'intérieur de la gaussienne dans la définition de la fonction de densité p_i vaut 0, ce qui résulte en un maximum global pour la fonction en question, et en un maximum local pour le diagramme 3D. Cela va nous permettre de décider à laquelle des 3 classes un point arbitraire a le plus de chance d'appartenir.



Contour plot des fonctions de densité



Graphe 3D des fonctions de densité

5. Visualisation des frontières

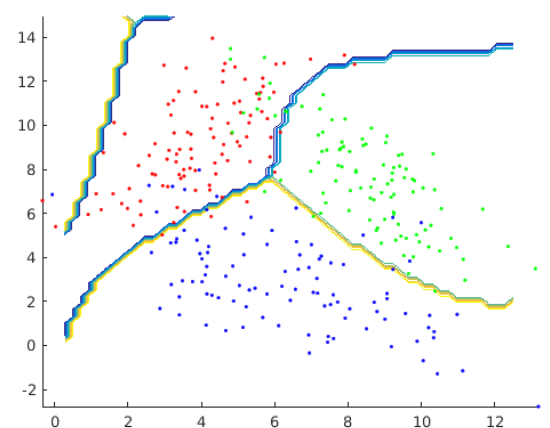
1) D'après la formule de Bayes, on sait que l'on doit choisir une classe d'appartenance ω_i pour x telle qu'elle maximise $p_i(x | \omega_i)$. Ainsi, pour chaque point x , nous n'avons qu'à stocker les trois densités de probabilité dans un tableau et utiliser comme classe l'index du maximum.

$$[a, Z(i, j)] = \max([dens1(i, j) \quad dens2(i, j) \quad dens3(i, j)]);$$

2) Le *contour plot* de Z nous donne l'image ci-contre. On voit clairement quelle région délimite quelle classe, mais il reste une subtilité : la région la plus à gauche (à gauche encore de la région rouge) dénote en fait **également** la classe 3 (bleue). Cela est dû à la forte variance du vecteur gaussien originel qui, comme le *contour plot* des fonctions de densité ci-dessus le montre, a une portée bien plus grande que les deux autres classes.

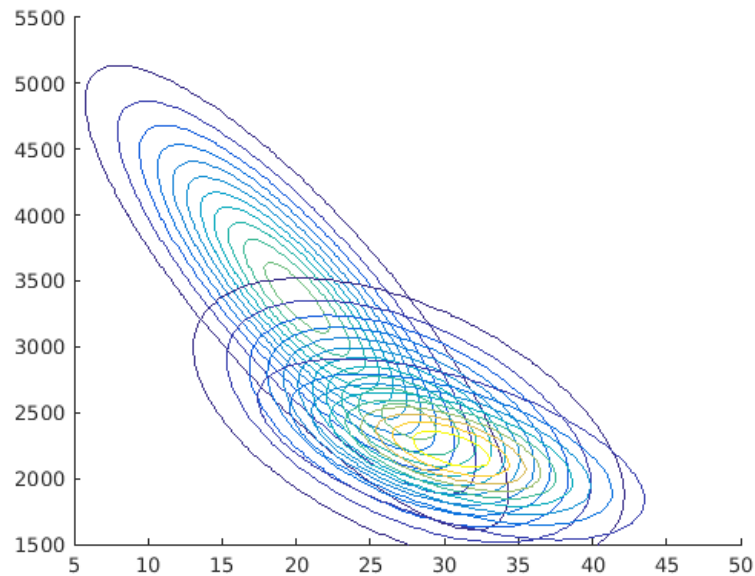
6. Application

L'application du même algorithme à un ensemble de voitures restreint à deux paramètres ne permet malheureusement pas de déterminer la classe d'appartenance de chaque véhicule. En effet, et comme mit en évidence par les images ci-dessous,

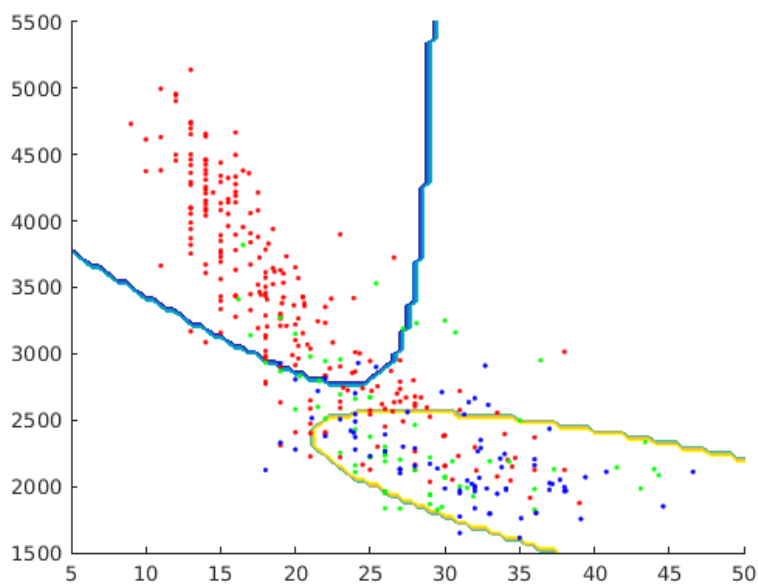


Superposition du contour plot de Z et des échantillons

beaucoup de points de classes différentes se retrouvent dans une même zone dans l'espace des paramètres, ce qui se traduit par des fonctions de densité qui se chevauchent et rend l'algorithme inefficace. Une solution serait de considérer plus de paramètres (soit plus de degrés de liberté) et de lancer le même algorithme agissant en dimension supérieure pour pouvoir plus certainement classer les échantillons.



Fonctions de densité des 3 classes



Contour plot de la détermination de classe