1. Régression linéaire et non linéaire régularisée
   1. **Régression linéaire:**
2. L’ensemble θ des paramètres est {w, b} : w étant le vecteur de poids de dimension Rd et b, le biais de dimension R.
3. Le risque empirique est
4. Pour minimiser le risque empirique, on cherche le θ qui donne le moins d’erreur sur l’ensemble d’entraînement, soit :
5. Le gradient du risque empirique est :

* 1. **Régression linéaire régularisée (“ridge regression”):**

1. Le gradient du risque régularisé est :

Expliquer la différence avec le risque non régularisé

1. DescenteDeGradientBatch(

←

faire

←

jusqu’à | | <

* 1. **Régression avec un pré-traitement non-linéaire fixe:**

1. k est de dimension 1
2. Avec x en dimension d = 2, on a:
3. Avec x en dimension d, on à de dimension

Partie 2: Programmation

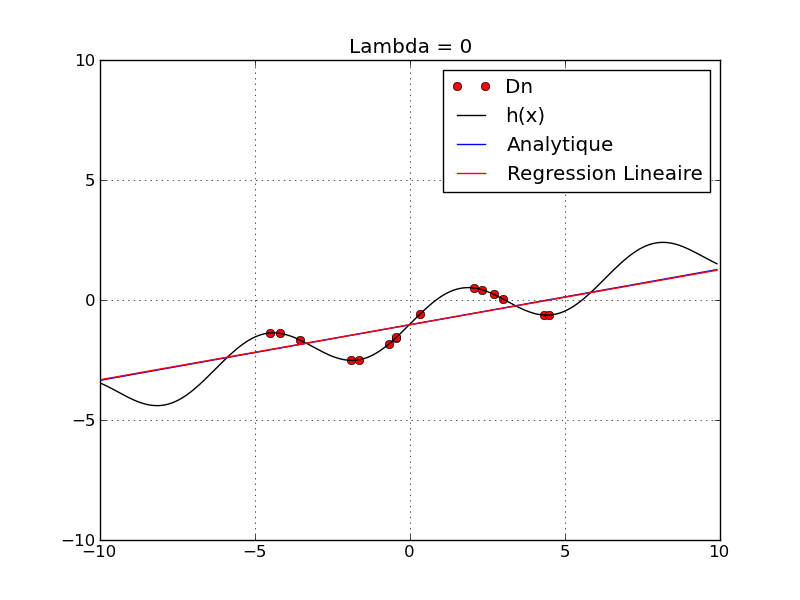
4)

Lambda = 0

Paramètres :

Analytique : w = 0.23179125, b = -1.02987640

Régression: w = 0.22932296, b = -1.03025983

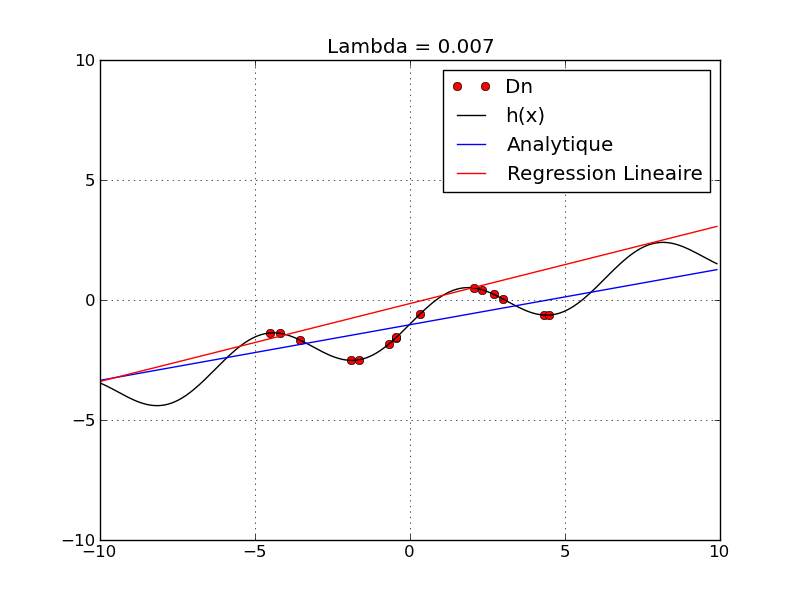


5)

Valeurs numériques pour un lambda (valeur extrême) = 0.007

Analytique : w = 0.23177799, b = -1.02987477

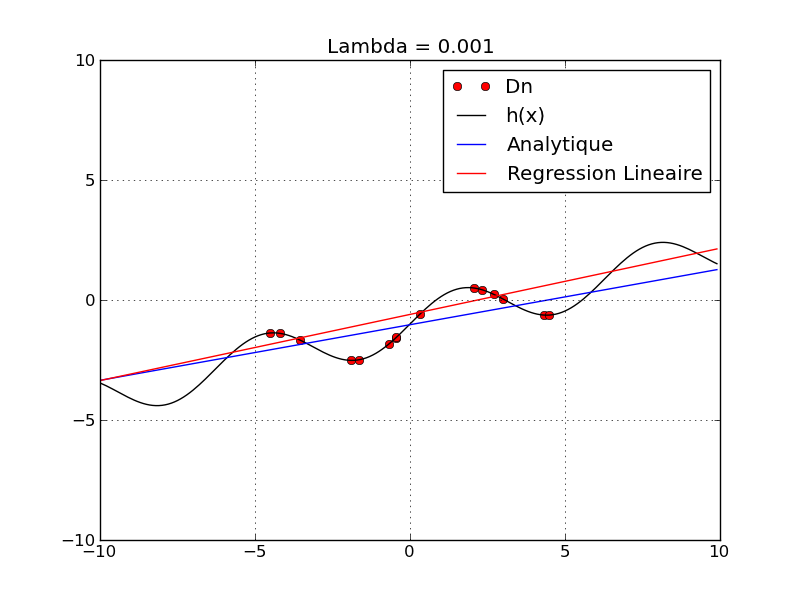
Régression: w = 0.32455948, b = -0.14699578



Valeurs numériques pour un lambda (valeur intermédiaire) = 0.001

Analytique : w = 0.23178936, b = -1.02987617

Régression: w = 0.27562477, b = -0.60097945

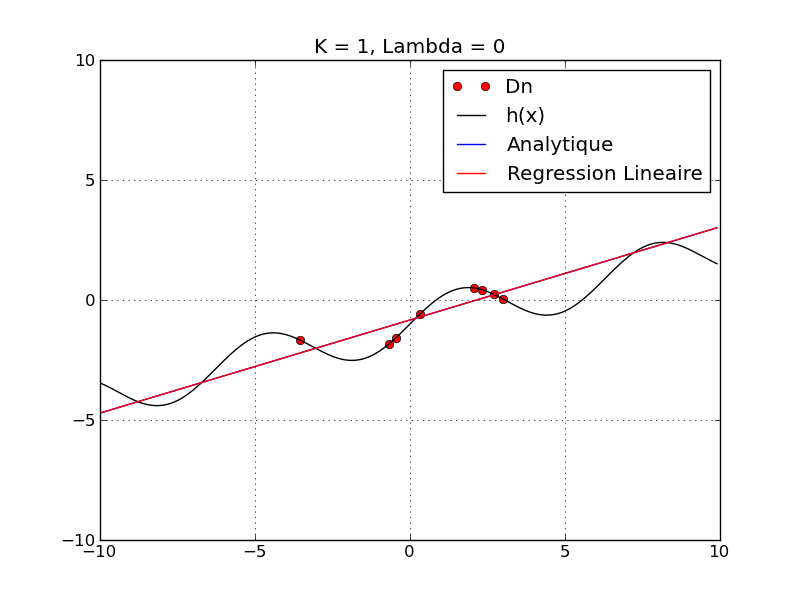


6)

k = 1

Regression Analytique : [-0.85256288 0.38943102]

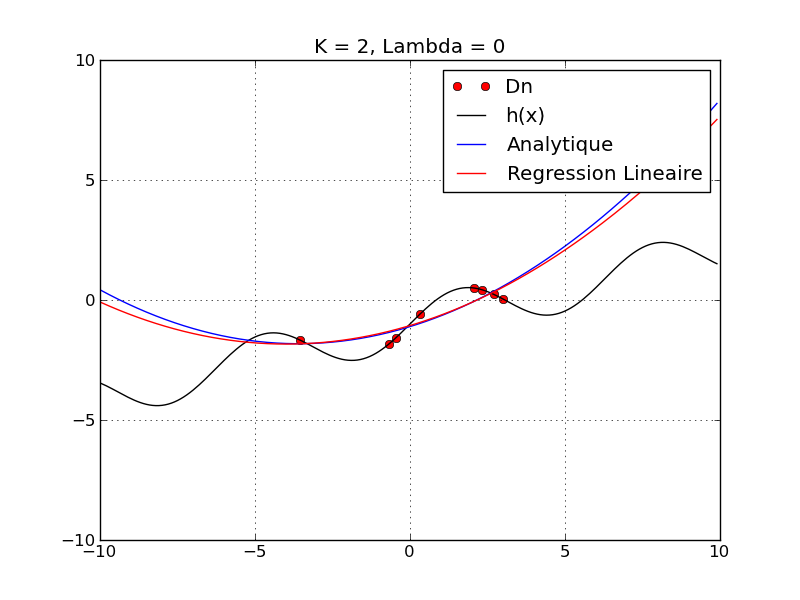
Regression Gradient : [ 0.38945179 -0.8505361 ]



k = 2

Regression Analytique : [-1.27421837 0.4044939 0.07332763]

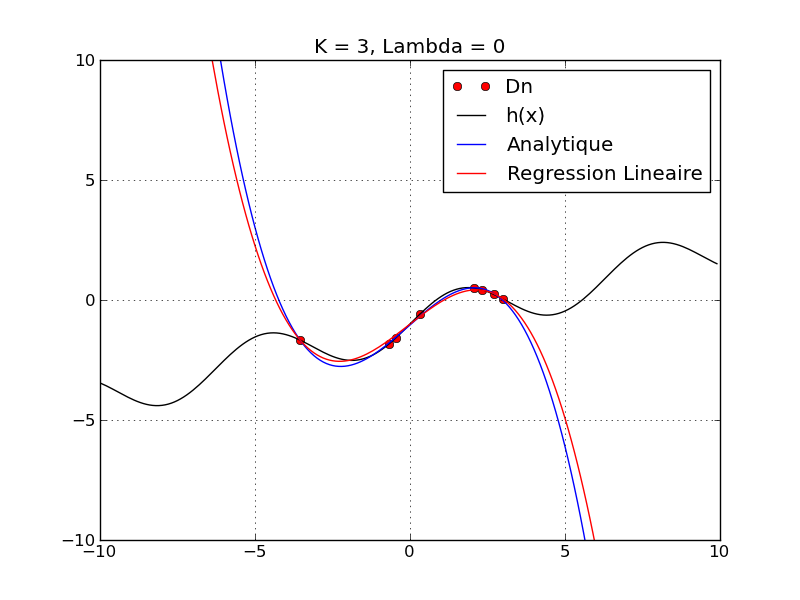
Regression Gradient : [ 0.38960116 0.0574411 -1.14491271]



k = 3

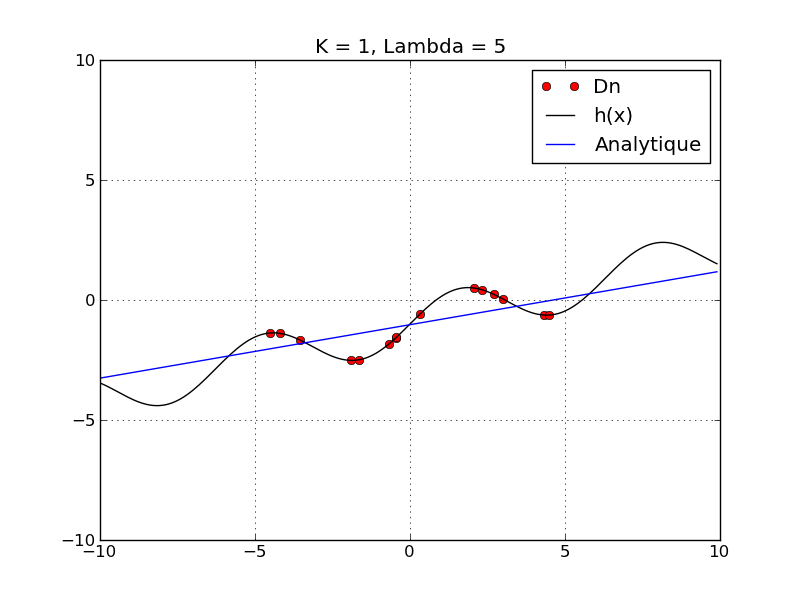
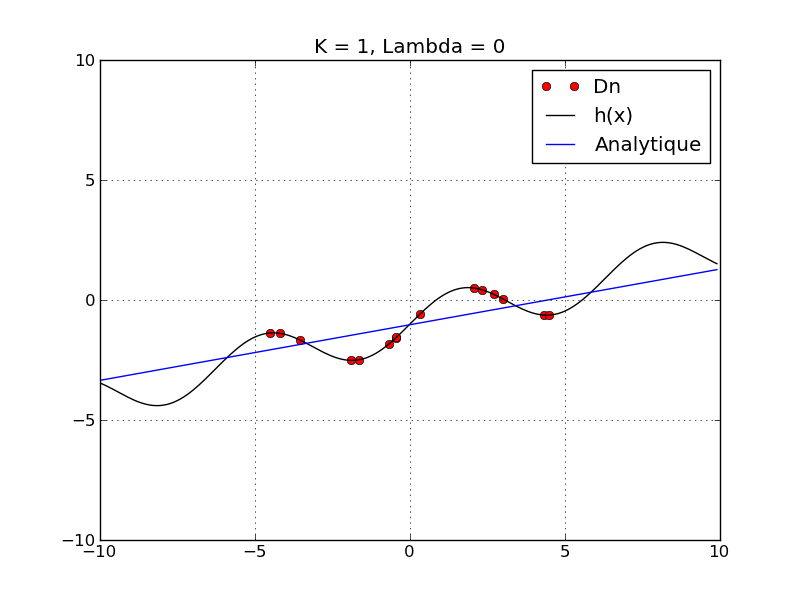
Regression Analytique : [-1.08185509 1.11511891 -0.01274914 -0.07913959]

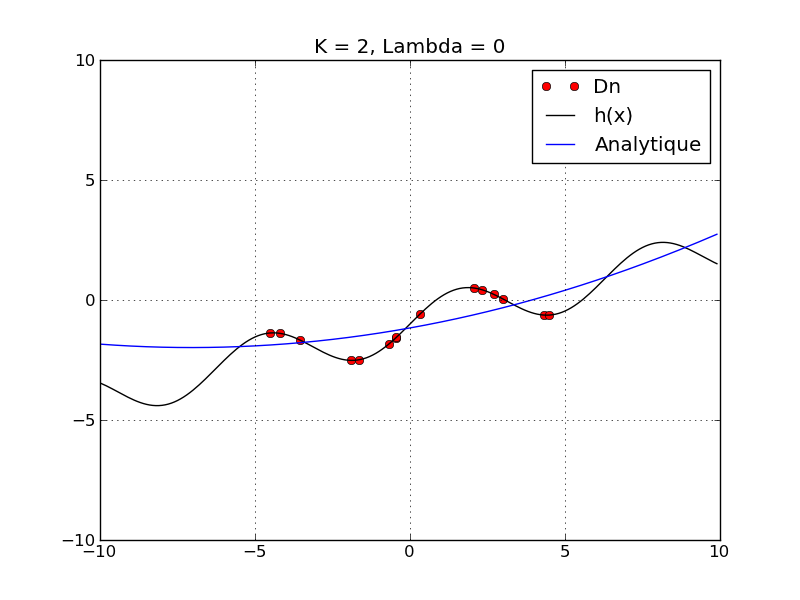
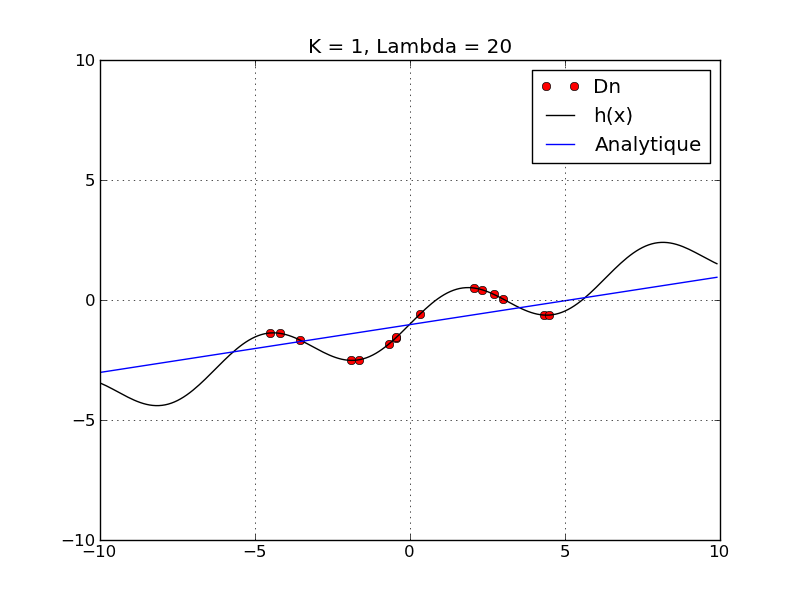
Regression Gradient : [ 1.07271974 -0.01828174 -0.07517304 -0.9904767 ]

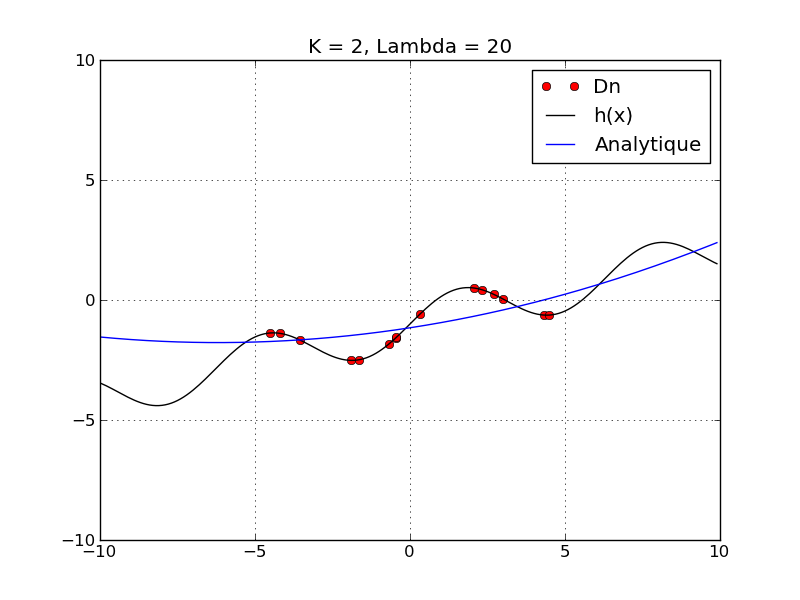
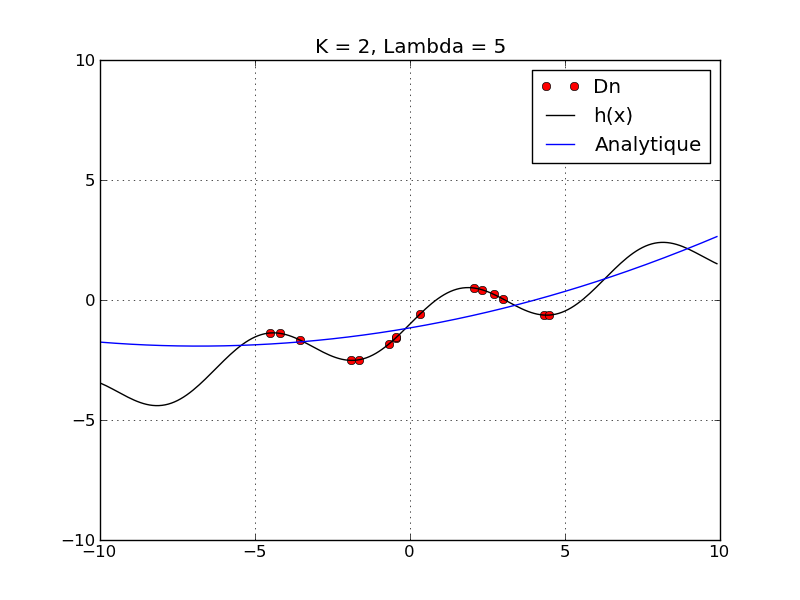


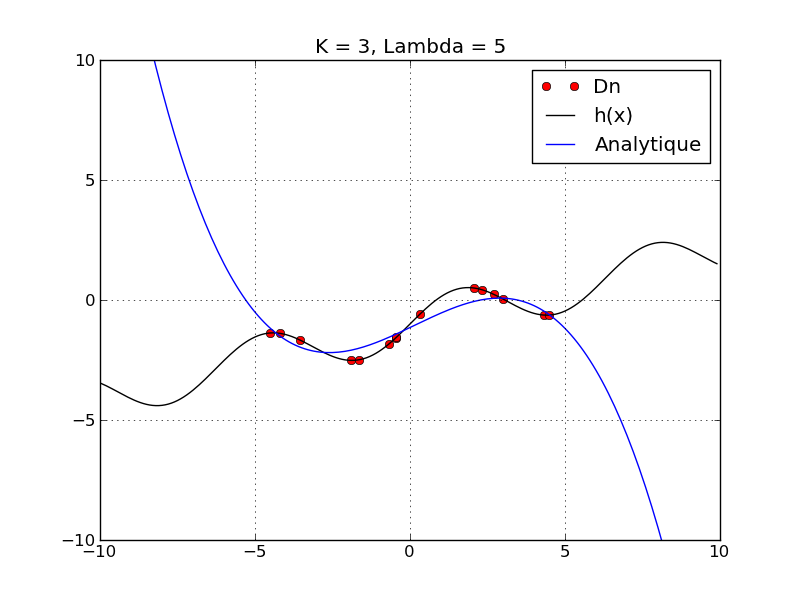
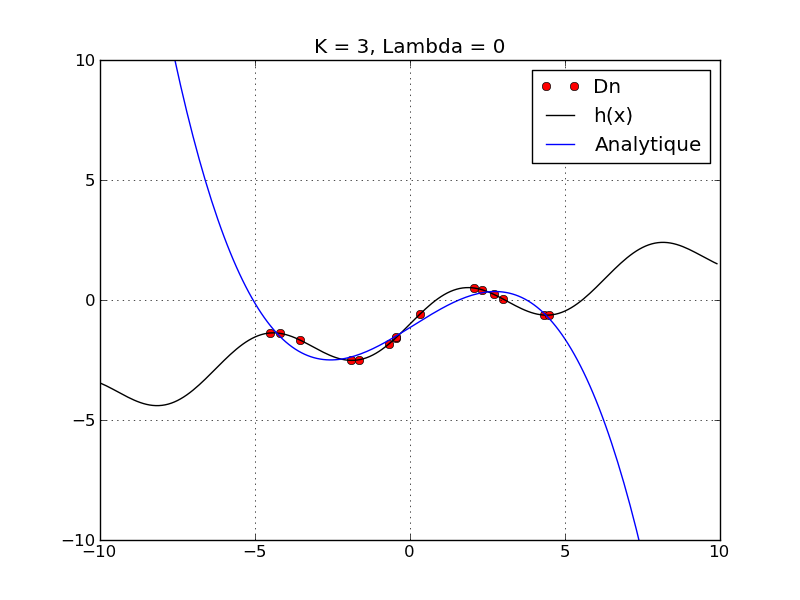
7) Effets de la variation de lambda pour un K donné (K = 1,2,3,10 ; Lambda = 0,5,20)

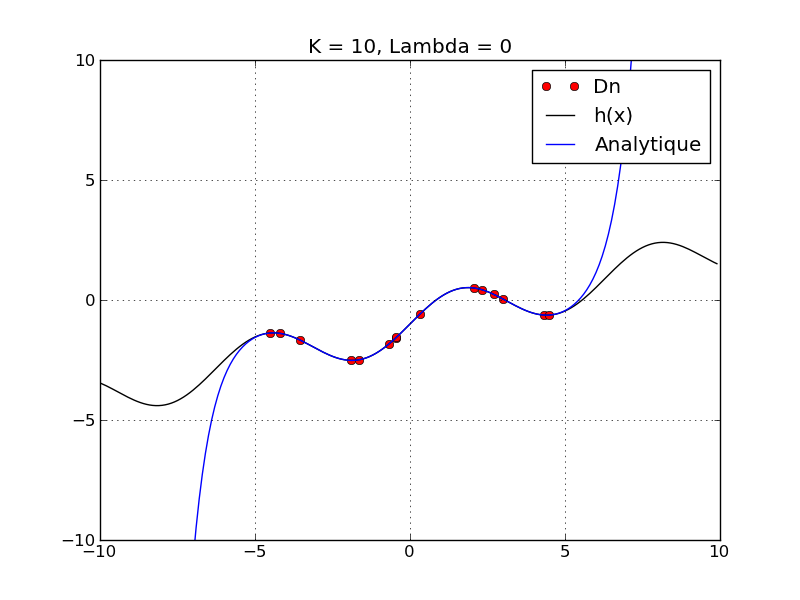
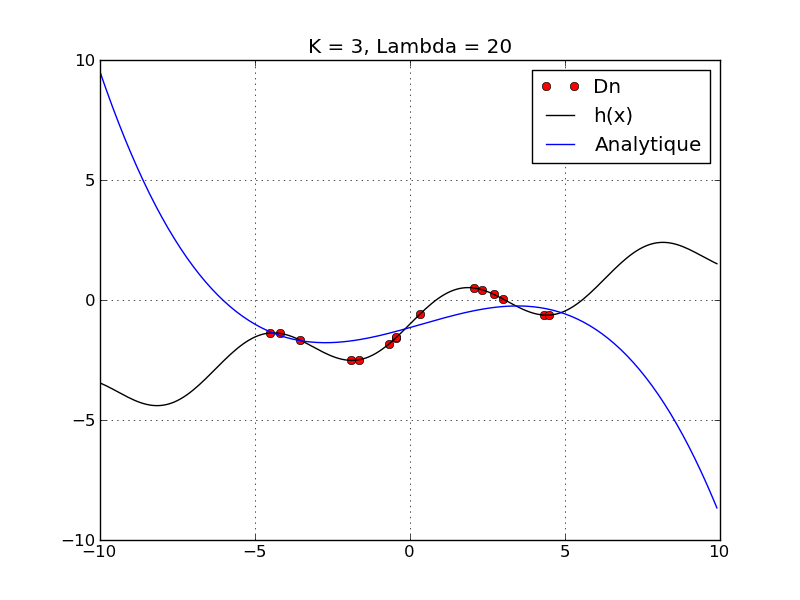
-> Expliquer l’influence de l’hyper-paramètre lambda

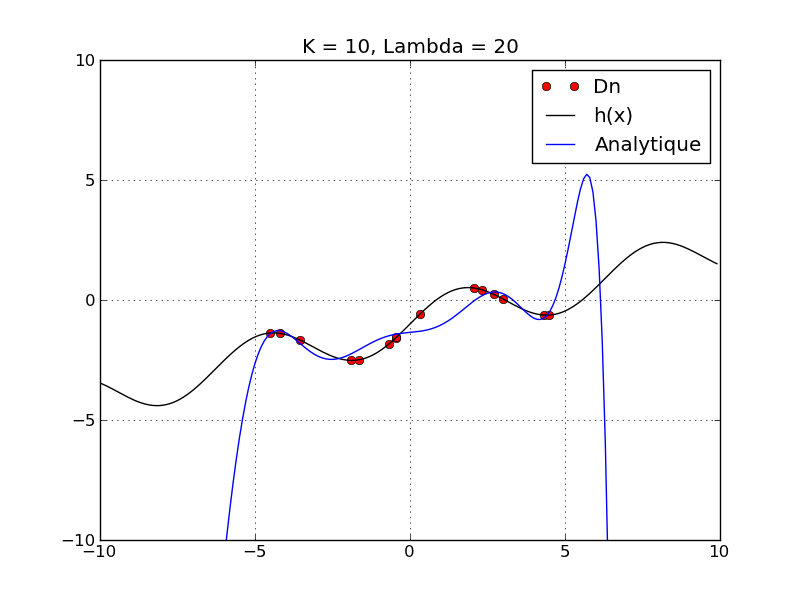
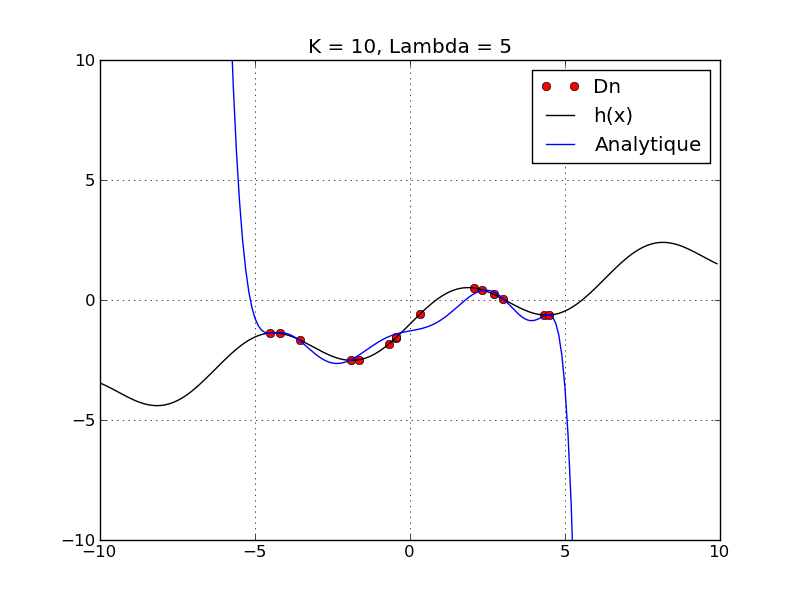












8) Nous appliquons maintenant l’algorithme de regression\_analytique sur le problème de classification 2D « ellipse.txt », avec des prétraitements k = 1, 2, 3, 4 pour des lambdas = 0, 5 , 20

