

Module - Apprentissage

Master Informatique, mention PRIM, 2005-2006

24 janvier 2006

1 Un robot qui a l'ouïe fine

Un robot se déplace dans un environnement bidimensionnel. Sa position dans un repère orthonormé donné est notée $(x, y) \in \mathbb{R}^2$. Dans l'environnement, on a placé trois hauts-parleurs h_1 , h_2 et h_3 , non alignés, dont on ne connaît pas la position. Ces hauts-parleurs émettent chacun en permanence un son sinusoïdal, de fréquence respective f_1 , f_2 et f_3 . L'intensité du son provenant d'un haut-parleur s'atténue au fur et à mesure que l'on s'éloigne de lui. Si d est la distance à la source, l'intensité du son est proportionnelle à $1/d^2$. Le robot est équipé d'un microphone standard¹ qui capte la résultante de ces trois sons, et un traitement du signal sonore fournit les intensités i_1 , i_2 et i_3 des sons des trois hauts parleurs perçus au point (x, y) où se situe le robot.

1.1 Localisation par apprentissage supervisé

On dote le robot d'un GPS qui lui fournit sa position (x, y) . L'idée est d'utiliser cette information pour associer les intensités enregistrées par le robot à sa position. Ainsi, après apprentissage, on pourra se passer de GPS, et calculer la position (x, y) du robot à partir de la mesure du triplet (i_1, i_2, i_3) .

1.1.1 Utilisation d'un perceptron mono-couche

Pour réaliser l'apprentissage supervisé, on choisit d'utiliser un perceptron mono-couche. On constitue une base d'exemple S de 200 échantillons.

Question 1.1 : Dessinez ce perceptron, et décrivez la base d'exemple (de quoi est-elle constituée?).

Question 1.2 : Que calcule le perceptron à partir de S ? Quelle hypothèse fait-on sur le problème, du moment que l'on considère pertinente l'utilisation du perceptron mono-couche?

Question 1.3 : Selon vous, cette hypothèse est elle légitime? Proposer une mesure qui, après apprentissage, confirme votre avis sur la légitimité de l'hypothèse. Vous argumenterez avec les concepts de théorie de l'apprentissage statistique vus en cours.

¹Non directionnel.

1.1.2 Utilisation d'une ϵ -SVM (régression)

On utilise maintenant une ϵ -SVM sur la même base S que précédemment. On vous rappelle qu'une ϵ -SVM est la SVM vue en cours pour calculer une régression. Si $S = (\xi_l, u_l)_{1 \leq l \leq n}$ sont les n éléments de la base d'exemple, cette SVM cherche (w, b) tel que « autant que faire se peut », $\forall l, |(w \cdot \xi_l + b) - u_l| < \epsilon$. Le processus d'optimisation de la SVM calcule un jeu de coefficients a_l et un b , à partir desquels est définie la fonction $f(\xi) = \sum_l a_l \cdot k(x_l, \xi) + b$. La fonction k est le noyau, et f ainsi calculé par la SVM est une interpolation de la base d'exemple (i.e. $f(\xi_l) \approx u_l$). Il existe une constante C qui paramètre l'algorithme, nous la fixerons à 1 dans tout le problème, sans remettre en cause ce choix. Dans les questions suivantes, on supposera que vous disposez d'une librairie pour résoudre les SVM, on ne vous demande donc jamais de décrire l'algorithme.

Question 1.4 : On choisit un noyau gaussien de paramètre σ , décrivez la mise en œuvre de l'apprentissage (précisez les données du problème qui jouent le rôle de ξ_l et u_l , etc...)

Question 1.5 : Si l'on choisit un noyau gaussien de paramètre σ trop petit, que se passe-t-il ? Et si σ est trop grand ? Vous argumenterez avec les concepts de théorie de l'apprentissage statistique vus en cours.

Question 1.6 : Comment trouver expérimentalement le bon σ ?

Question 1.7 : La qualité de prédiction de la position est-elle encore bien conservée si l'on est loin d'un point de mesure de S ? Justifiez.

1.2 Navigation à l'oreille

On souhaite appliquer un algorithme d'apprentissage par renforcement basé sur les MDP² pour faire naviguer notre robot. On n'a plus recours au GPS, et on ne s'appuiera pas sur les prédictions de (x, y) réalisées précédemment. En fait, on identifiera un lieu par le triplet (i_1, i_2, i_3) lui correspondant, sans chercher à retrouver (x, y) . Les intensités à la sortie des hauts-parleurs ont une valeur de 100, et on rappelle qu'elles s'atténuent en $1/d^2$. Comme les méthodes à base de MDP sont discrètes, on se propose de discrétiser l'espace des perceptions, puis dans un deuxième temps de résoudre un MDP pour naviguer.

1.2.1 Discrétisation de l'espace des perceptions

On définit le type `Son` tel que si `a` est de type `Son`, on a `a.i1`, `a.i2` et `a.i3` qui sont les 3 mesures d'intensité au point `a`. On se fixe pour représenter l'espace des perceptions du robot un tableau `S` de 1000 cases, chaque case `S[i]` étant de type `Son`³

On coupe la plage de variation de $[0, 100]$ en 10 intervalles réguliers ($v_0 = [0, 10[$, $v_1 = [10, 20[$, \dots , $v_8 = [80, 90[$ et $v_9 = [90, 100]$). Pour une intensité i , on note $v(i)$ le numéro de l'intervalle auquel elle appartient. Par exemple $v(18) = 1$, $v(35) = 3$, $v(91) = 9$. On remplit alors le tableau `S` de la façon suivante :

```
for centaine=0 to 9
  for dizaine=0 to 9
```

²Markovian Decision Processes.

³On a `S[i].i1`, `S[i].i2` et `S[i].i3`.

```

for unite=0 to 9
    etat = centaine*100 + dizaine*10 + unite
    S[etat].i1 = centaine*10+5
    S[etat].i2 = dizaine*10+5
    S[etat].i3 = unite*10+5

```

Ainsi, lorsque le robot reçoit une perception (i_1, i_2, i_3) , on dira qu'il est dans l'état $e = 100v(i_1) + 10v(i_2) + v(i_3)$, et on considérera cette perception comme étant $S[e]$.

Question 1.8 : Expliquez en quoi cette discrétisation utilise mal la place mémoire occupée par S .

Question 1.9 : Montrez que la distribution des (i_1, i_2, i_3) dans $[0, 100]^3$ est une surface.

Question 1.10 : Déduisez des réponses aux deux questions précédentes un algorithme de quantification vectorielle qui, à partir d'une collection de perceptions reçues par le robot, permet d'exploiter mieux la mémoire S .

Question 1.11 : Que se passe-t-il si dans la collection utilisée, beaucoup de perceptions ont été acquises avec un robot à l'arrêt en un point ? Comment nettoyer la collection pour éviter ces effets ?

1.2.2 Construction d'un MDP

Le robot dispose de 3 actions :

1. Avancer d'une distance d ,
2. tourner sur place de 10° puis avancer de d ,
3. tourner sur place de -10° puis avancer de d .

On fixe au robot un objectif sous la forme d'un état perceptif (i_1, i_2, i_3) à atteindre, et l'on souhaite mettre en œuvre un algorithme de renforcement pour que le robot apprenne à utiliser ses perceptions pour atteindre l'objectif, en déclenchant une suite d'actions.

Quand le robot perçoit (i_1, i_2, i_3) , l'état correspondant est la case de $S[e]$ dont les valeurs sont les plus proches de la perception.

Question 1.12 : Pourquoi ne peut-on pas utiliser S comme espace d'état ?

Pour résoudre ce problème, on fixe un deuxième micro sur le robot. Les deux micros sont fixés de façon rigide sur le robot, de sorte qu'ils soient espacés d'une distance d . Des deux micros, on extrait comme précédemment deux cases $S[e1]$ et $S[e2]$.

Question 1.13 : Justifiez que les couples $(e1, e2)$ forment un espace d'états utilisable pour un MDP.

Après une phase d'apprentissage, on souhaite pouvoir poser un objectif (i_1, i_2, i_3) quelconque et laisser le robot s'y rendre, sur la base de ce qu'il a appris.

Question 1.14 : L'apprentissage peut-il être réalisé par un Q-Learning ? Justifiez

Question 1.15 : Si vous avez répondu oui à la question précédente, décrivez la mise en œuvre du Q-Learning, si vous avez répondu non, proposer une autre méthode qui soit adaptée.