Master RAR
15 Janvier 2013
9-12h

Reconnaissance des formes

Epreuve correspondant au cours de Marie-Odile Berger
Documents distribués en cours autorisés
Rédiger les parties de Marie-Odile Berger et Bernard Girau sur copies séparées.
J'attends de vous des réponses précises et argumentées aux questions posées.

1 Estimation

On considère un ensemble de n points 2D $(x_i, y_i)_{1 \leq i \leq n}$, par exemple comme celui indiqué dans la figure 1. On veut approcher ces points par une droite. On choisit d'estimer les paramètres de la droite avec les critères suivantes:

1. $min_{a,b} \sum_i (y_i - ax_i - b)^2$
2. $min_{a,b} \sum_| y_i - ax_i - b|$
3. $min_{a,b,c} \sum \frac{|ax_i+by_i+c|}{\sqrt(a^2+b^2)}$

Figure 1: Un exemple de données dont il faut extraire la droite et calculer au mieux son équation.

Expliquer l'intérêt les défauts de ces trois critères notamment vis à vis des problèmes suivants: (i) présence de points aberrants, (ii) invariance du résultat vis à vis d'un changements du repère d'acquisition des données, (iii) facilité de calcul des paramètres de la droite.

2 Prédicteur linéaire

On considère le problème d'estimer une grandeur à partir de mesures. Pour fixer les idées, on considère un robot muni de capteurs se déplaçant dans le plan. L'objectif est de pouvoir calculer le déplacement t du robot à partir du vecteur des mesures $X, (X \in R^k)$ fournies par les capteurs. La façon de calculer t à partir du vecteur de mesures X n'est pas connue. On cherche dans cet exercice à construire un prédicteur linéaire H permettant de déduire la position t du vecteur d' observations X par une formule de type HX, H étant une matrice. L'objectif est d'apprendre la matrice H à partir d'exemples.

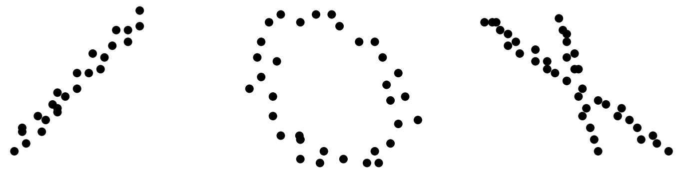
1. Apprentissage de H
   On génère un ensemble de m données d'apprentissage $A = \{(t_i, X_i)\}_{1 \leq i \leq m}$, ou t_i est le déplacement effectué et X_i le vecteur contenant les mesures des capteurs. Ces données

d'apprentissage sont par exemple acquises en faisant effectuer des mouvement au robot muni d'un GPS, ce qui permet de mesurer directement t. On note $T = [t_1, .., t_m]$, la matrice $2 \times m$ contenant les positions apprises et $X = [X_1, ..., X_m]$ la matrice $k \times m$ contenant les vecteurs de capteurs mesurés. On cherche donc à estimer H au mieux respectant la relation $T = HX$. Montrer que l'estimation de H aux moindres carrés est donnée par $H = TX^t(XX^t)^{-1}$.

2. Le modèle linéaire a été choisi a priori. Comment peut on valider (où non) que ce modèle est bien adapté au problème?

3 ACP et ICA

1. La figure ci-dessous montre trois exemples de données dans le plan. Expliquer dans chacun des cas si une analyse en composantes principales (ACP) peut correctement décrire ces données. Justifiez votre réponse.



(a) (b) (c)

2. Une analyse en composantes indépendantes (ICA) fonctionnerait-t-elle mieux sur ces données? Pourquoi?

4 Question diverses

- Le point de rupture d'un estimateur est la proportion d'observations incorrectes qu'un estimateur peut supporter sans donner de valeurs très erronées Expliquer pourquoi le point de rupture de l'estimateur moyenne est 0. Quel est le point de rupture de l'estimateur médiane?

  On considère l'estimateur suivant d'une grandeur p: $\hat{p}_h = \sum_{i=1}^{h} r_i(p)$ où les résidus $r_i(p)$ sont ordonnés par valeurs croissantes (pour fixer les idées, $p = (a, b)$ est une droite à estimer et $r_i(p) = (y_i - ax_i + b)^2$ est le résidu associé à chaque mesure (x_i, y_i). Quel est le point de rupture de cet estimateur?

- On considère une variable aléatoire X décrivant l'état d'un système. Soit Λ la matrice de covariance associée. Cette variable subit une transformation f l'amenant à une position $Y = f(X)$. On considère les deux cas suivants:

  1. f est une fonction lisse
  2. f est non dérivable

  Décrivez dans les grandes lignes dans chacun des ces deux cas la méthode qui vous semble la mieux adaptée pour calculer ou approximer la matrice de covariance de $Y = f(X)$. Vous préciserez les performances de la méthode choisie en terme de rapidité de calcul et de fidélité par rapport à la valeur théorique de Y.

- On considère le problème d'estimation aux moindres carré d'une grandeur θ à partir de mesures Z liés par l'équation $Z = A\theta$. On choisit pour métrique une norme significative du problème, $||Z||^2 = Z^t\Lambda^{-1}Z$ où Λ est la matrice de covariance des mesures.

  – Quel est l'intérêt de choisir une telle métrique?
  – montrer que la solution de l'équation $Z = A\theta$ aux moindres carrés est donnée par

$$\hat{\theta} = (A^t\Lambda^{-1}A)^{-1}A^t\Lambda^{-1}Z$$