DIRO
IFT 2425

DEVOIR N° 1

Bissection/Point Fixe & Monstres Numériques

Max Mignotte

DIRO, Département d'Informatique et de Recherche Opérationelle.
http : //www.iro.umontreal.ca/∼mignotte/IFT2425/
E-mail : mignotte@iro.umontreal.ca

I. Méthodes Dichotomique et du Point Fixe

Introduction

Le principal intérêt de la loi de distribution de Weibull réside dans sa flexibilité. Ainsi, pour certaines valeurs d'un de ses paramètres (le paramètre de forme c), on retrouve la loi exponentielle ($c = 1$) et la loi de Rayleigh ($c = 2$). Cette propriété explique l'utilisation qu'en font les statisticiens pour modéliser des échantillons de variables aléatoires exponentielles dont les critères de dépendances ne sont pas pleinement vérifiés.

En traitement d'image, cette flexibilité est aussi particulièrement intéressante pour modéliser la distribution du bruit de speckle dans une image SONAR [1], ou des niveaux de gris d'une images RADAR de la surface de la mer, ou encore pour modéliser la norme du gradient des niveaux de gris associé aux contours d'une image [2]. La loi de distribution de Weibull est définie par :

$$\mathcal{W}(y; c, \alpha) = \frac{c}{\alpha}\Big(\frac{y}{\alpha}\Big)^{c-1} \cdot \exp\Big(-\frac{y^c}{\alpha^c}\Big)$$

avec $y > 0$, $\alpha > 0$ et $c > 0$. α et c sont respectivement un paramètre d'*échelle* et de *forme* strictement positifs. Nous avons représenté en Fig. 1 le tracé de cette loi pour différentes valeurs du facteur de forme c.

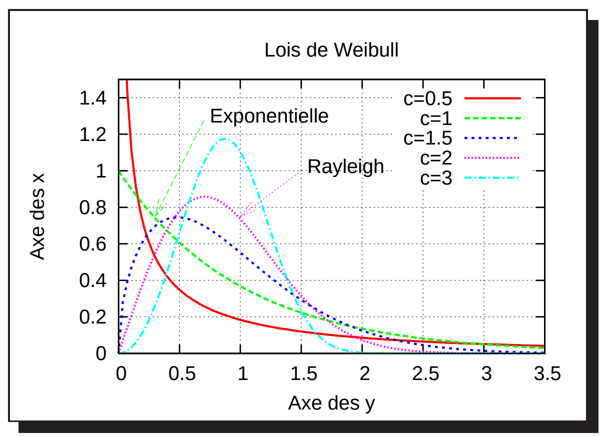

FIG. 1 – Lois de Weibull pour différentes valeurs du facteur de forme c (dans cet exemple, $\alpha = 1$)

La méthode du Maximum de Vraisemblance (MV) permet d'obtenir une estimation du vecteur de paramètres $\Phi_y = (\alpha, c)$ pour un échantillon y donné (c'est-à-dire pour un ensemble de N valeurs aléatoires y_i suivant la distribution de Weibull). Si on suppose l'indépendance des différentes variables aléatoires y_i de l'échantillon y, la fonction de vraisemblance $\mathcal{L}$ s'écrit

$$\mathcal{L}(\Phi_y) = P(y/\Phi_y) = \Big(\frac{c}{\alpha^c}\Big)^N \prod_{i=1}^{N} \Big[y_i^{(c-1)} \exp\Big(-\frac{y_i^c}{\alpha^c}\Big)\Big]$$

Pour déterminer $\hat{\alpha}_{\text{MV}}$ et $\hat{c}_{\text{MV}}$ (valeurs qui maximisent la fonction de vraisemblance), on doit résoudre simultanément les deux équations suivantes $\partial \ln \mathcal{L}(\Phi_y)/\partial\alpha = 0$ et $\partial \ln \mathcal{L}(\Phi_y)/\partial c = 0\,,$

conduisant au système d'équations

$$\begin{aligned} -N + \frac{1}{\hat{\alpha}^{\hat{c}}} \sum_{i=1}^{N} y_i^{\hat{c}} &= 0 \\ \frac{N}{\hat{c}} + \sum_{i=1}^{N} \ln y_i - N \ln \hat{\alpha} - \sum_{i=1}^{N} \Big(\frac{y_i}{\hat{\alpha}}\Big)^{\hat{c}} \ln\Big(\frac{y_i}{\hat{\alpha}}\Big) &= 0 \end{aligned}$$

Après simplification, on trouve une forme explicite pour $\alpha_{\rm MV}$: $\alpha_{\rm MV} = \big(\frac{1}{N}\sum_{i=1}^{N} y_i^{c_{\rm MV}}\big)^{\frac{1}{c_{\rm MV}}}$. Par contre on ne peut pas obtenir de forme explicite pour l'estimateur $c_{\rm MV}$. En effet $c_{\rm MV}$ est solution de l'équation

$$f(c_{\rm MV}; y_i) = \frac{\sum_{i=1}^{N}\big(y_i^{c_{\rm MV}} \ln y_i\big)}{\sum_{i=1}^{N} y_i^{c_{\rm MV}}} - \frac{1}{c_{\rm MV}} - \frac{1}{N}\sum_{i=1}^{N} \ln y_i = 0 \qquad (1)$$

qui n'a pas de solution analytique explicite et qui ne peut se résoudre que numériquement (par approximations successives) par l'une des méthodes numériques vues en cours [1,2].

Tab. 1 – Valeurs des $N = 10$ points de données y_i

y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}
0.11	0.24	0.27	0.52	1.13	1.54	1.71	1.84	1.92	2.01

Méthode Dichotomique

Puisque l'estimateur du MV est unique, nous savons donc que l'Équation (1) admet forcément une solution. De plus, comme on sait que $c_{\rm MV} > 0$, il suffira donc de démarrer la méthode de la *bissection* avec un intervalle dont la borne inférieure sera une valeur presque nulle ($c_{\rm MV}^{\rm min} = 0.01$) et la borne supérieure une valeur (que l'on sait être) trop grande pour $c_{\rm MV}$ ($c_{\rm MV}^{\rm max} = 10$, par exemple)[1].

1– Programmer le méthode de bissection pour la recherche de la racine $\hat{c}_{\rm MV}$ de l'Équation (1) en utilisant des nombres flottants, l'intervalle $[c_{\rm MV}^{\rm min}, \ldots, c_{\rm MV}^{\rm max}]$ et avec comme critère d'arrêt, une tolérance sur $\hat{c}_{\rm MV}$ et une tolérance sur f de 10^{-6}.

Méthode du Point Fixe

2– Le raisonement précédent nous a permis de montrer qu'il existait une solution unique $\hat{c}_{\rm MV}$. Trouver maintenant expérimentalement[2] une forme adéquate transformant $f(c_{\rm MV}; y_i) = 0$ en une relation du type $c_{\rm MV} = g(c_{\rm MV}; y_i)$ puis en une suite itérative de la forme $c_{\rm MV}^{[n+1]} = g(c_{\rm MV}^{[n]}; y_i)$ qui converge vers la solution lorsque l'estimée initiale (premier terme de la suite) est $c_{\rm MV}^{[0]} = 0.5$, avec comme critère d'arrêt, une tolérance sur $\hat{c}_{\rm MV}$ et une tolérance sur f de 10^{-6}.

[1] Notons que ce petit raisonement, propre à ce problème de recherche d'estimateur MV, nous permet de savoir que la méthode de la *bissection* va nécesssairement converger vers une solution unique. Dans le cas où un tel raisonement n'est pas possible, on rappelle que l'on doit vérifier (avant d'appliquer la méthode de la *bissection*) que l'intervalle de départ encadre bien la solution et que la fonction $f(c_{\rm MV}; y_i)$ est continue et monotone sur cet intervalle de départ.

[2] En toute rigueur (i.e., théoriquement et non plus expérimentalement comme dans ce TP), on rappelle que l'on devrait montrer (cf. cours) que la fonction $g(x)$ choisie dans la relation itérative $c^{[n+1]} = g(c^{[n]})$ est telle que $|g'(x)| < 1, \ \ \forall x \in J$, un intervalle comprenant la racine et le premier élément de la suite $c^{[0]}$.

II. Monstre Numérique

Le terme *Monstre Numérique* vient de l'excellent article d'Essex *et al.* [3] répertoriant quelques célèbres exemples de problèmes d'erreures numériques existants en notation flottantes. Un exemple bien connu est la fonction

$$y(x) = (x^2 - 2x + 1) - (x - 1)^2$$

Une arithmétique idéale donnerait bien sur $y(x) = 0$ quelque soit la valeur de x. Mais un calcul en notation flotante fait apparaître la soustraction de deux nombres approximés quasi-égaux qui peut s'avérer être une catastrophique numérique (cf. cours) et ceci restera vrais tant que les compilateurs ne seront assez intelligent pour faire un ré-arrangement algébrique automatique de cette formule !

En utilisant, un programme et compilateur c++ et en affichant la courbe $y(x)$ lorsque x évolue dans l'intervalle $1.0 \pm 7.0 \times 10E-5$ et y dans l'intervalle $0.0 \pm 0.8 \times 10E-8$, un intéressant graphe (cf. Fig. 2) émerge qui nous montre à quoi ressemble le motif généré par ce monstre numérique ...

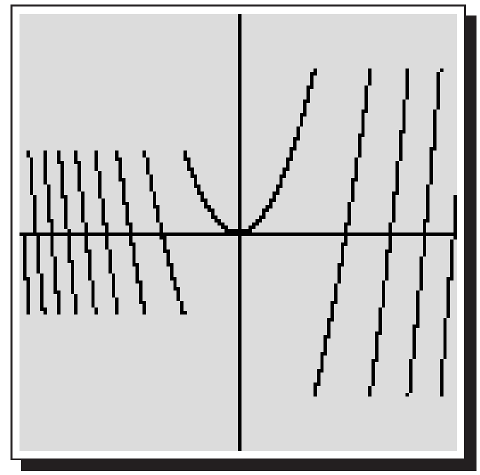

Fig. 2 – Motif généré par le monstre numérique $(x^2 - 2x + 1) - (x - 1)^2$ sur $x = 1.0 \pm 7.0 \times 10E-5$ et $y = 0.0 \pm 0.8 \times 10E-8$

3– Retrouver ce graphe en affichant intelligement le résultat $y(x)$ (pour x et y évoluant dans le bon intervalle) dans un tableau deux dimensions qui s'affichera ensuite à l'écran sous forme d'image[3].

Réferences

[1] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy. Three-class Markovian segmentation of high resolution sonar images. *Computer Vision and Image Understanding*, 76(3) :191–204, 1999.

[2] F. Destrempes and M. Mignotte. A statistical model for contours in images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26(5) :626–638, 2004.

[3] C. Essex and M. Davison and C. Schulsky. Numerical Monsters. *SIGSAM Journal*, 34(4) :16-32, December 2000.

[1]

Remise & Rapport

Vous devez rendre électroniquement le(s) programme(s) fait en C avant la date de remise spécifiée dans le fichier *barême* situé sur la page web du cours. Pour la remise électronique, utilisez le programme *remise* (*man remise* pour plus de détails) pour remettre votre code dans le répertoire TP<Numéro du Tp>. N'oubliez pas d'inscrire vos noms, courrier électronique en commentaire en haut du fichier .c remis. Les noms des programmes à remettre devront avoir le format suivant *Tp<Numero du Tp>-IFT2425-<Numéro de la question>.c*. Les programmes devront se compiler et s'executer sur Linux tel qu'indiqué dans le barême.

[3]

Conseils Pratiques

Utiliser le programme que je vous donne sur ma page web en initialisant la variable FLAG_GRAPH=1 et remplisser le tableau 2D nommé Graph2D dans le programme (initialement remplie de blanc) comme une feuille de papier dessin ou vous tracerez le graphe dans l'intervalle considéré avec le centre de l'axe des abscisses et des ordonnées au centre de ce tableau 2D. En utilisant la variable FLAG_GRAPH=1 , le programme que je vous donne sauvera cette image au format PGM. Vous pourrez ensuite la lire avec les logiciels DISPLAY, XV ou GIMP disponible sur Linux.
Attention, si vous changez de compilateur, il faudra sans doute changer les bornes de l'intervalles de x et y car les compilateurs gèrent les erreurs et les flottants de façon différente !