

TP 1 : LA SEGMENTATION BAYESIENNE

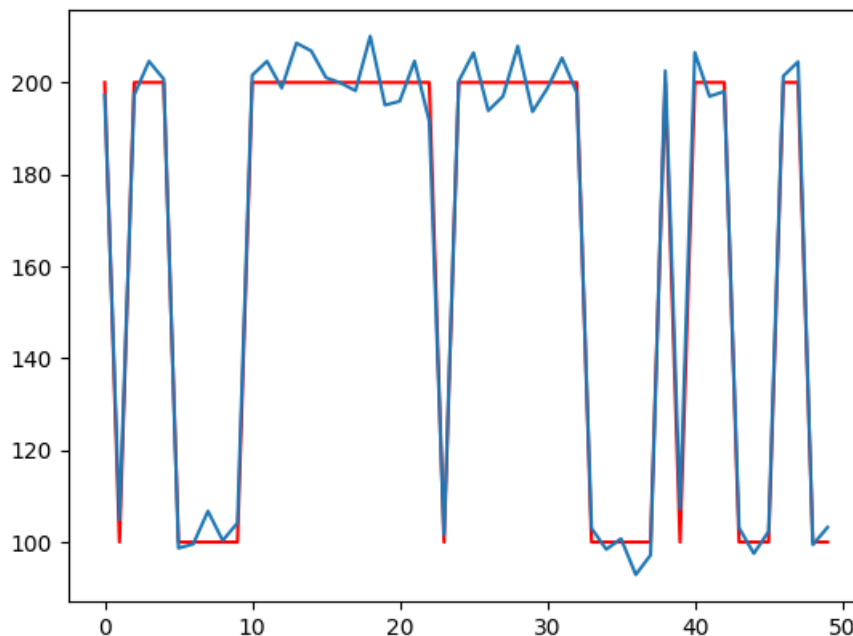
Par Victor LEDEZ (ledezvictor@gmail.com, victor.ledez@telecom-paris)

Lien du code sur GitHub : <https://github.com/VictorLEDEZ/tp-1>

1 UNE PREMIERE IDEE DES ENJEUX DU PROBLEME

1.A. FONCTION QUI BRUITE LE SIGNAL Y.

- Courbe rouge = le signal d'entrée
- Courbe bleu = le signal bruité (signal de réception)



Le bruit a été simulé avec les paramètres suivants :

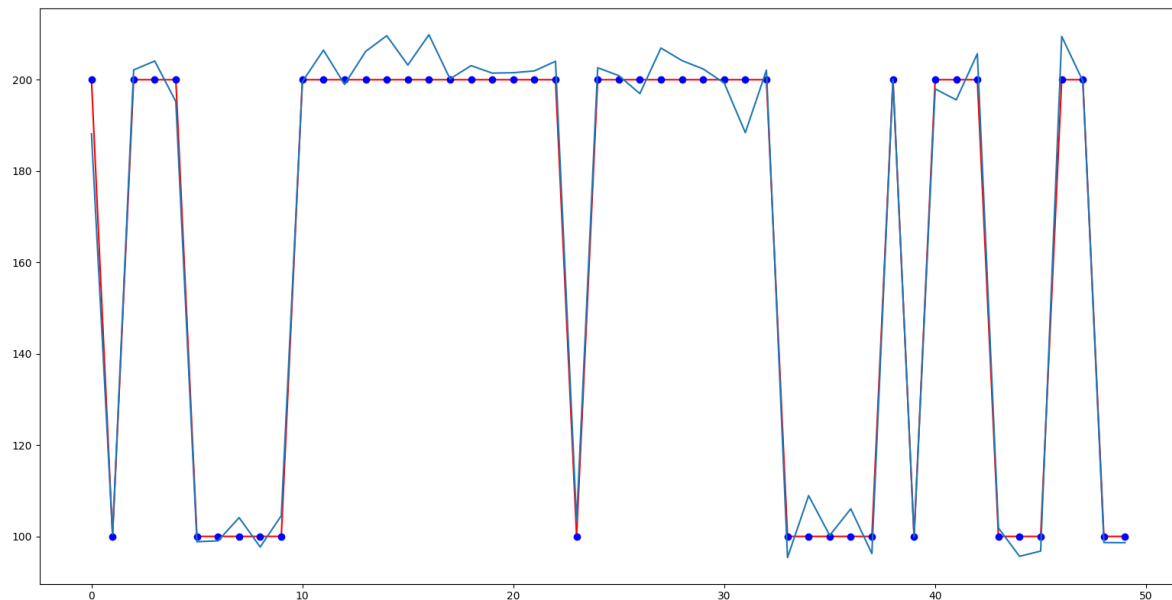
- $\mu_1 = 100$
- $\sigma_1 = 4$
- $\mu_2 = 200$
- $\sigma_2 = 5$

1.C. SUR UN MEME GRAPHIQUE LES COURBES DU SIGNAL ORIGINAL, DU SIGNAL BRUITE ET DU SIGNAL SEGMENTE.

En rouge = le signal d'entrée

En bleu = le signal bruité

En pointillés bleus = le signal reconstitué



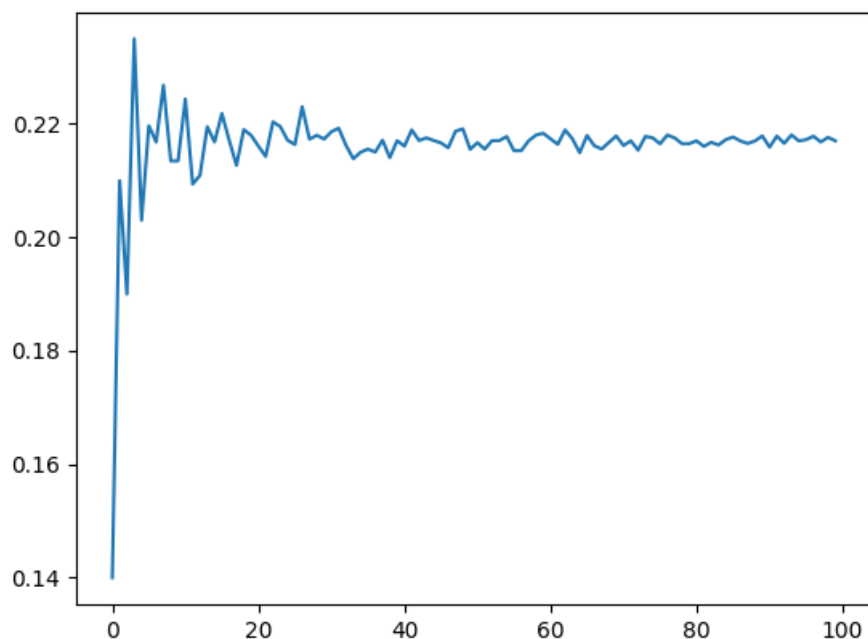
Le bruit a été simulé avec les paramètres suivants :

- $\mu_1 = 100$
- $\sigma_1 = 4$
- $\mu_2 = 200$
- $\sigma_2 = 5$

Nous constatons que la reconstitution a marché à 100% puisque le signal segmenté a été reconstruit en étant le même que le signal d'entrée. Ce qui est prévisible puisqu'il y avait un faible niveau de bruit.

3. TAUX D'ERREURS MOYEN

Voici le graphique représentant le taux d'erreur moyen au bout de 100 simulations :



Le bruit a été simulé avec les paramètres suivants :

- $\mu_1 = 127$

- $\text{Sigma1} = 1$
- $\text{Mu2} = 128$
- $\text{Sigma2} = 1$

Nous remarquons que le taux d'erreur moyen converge vers une valeur (ici 0.22 soit 22% d'erreur en moyenne sur le signal reconstitué).

Ce phénomène s'explique par la théorie des grands nombres ; plus on répète l'expérience, plus les résultats convergent vers les mêmes nombres. Ici, après chaque itération, le taux moyen est calculé en refaisant la moyenne avec les taux précédents donc, le taux converge puisque chaque taux a à peu près la même valeur.

En matière de niveau de bruit nous constatons que celui-ci se stabilise.

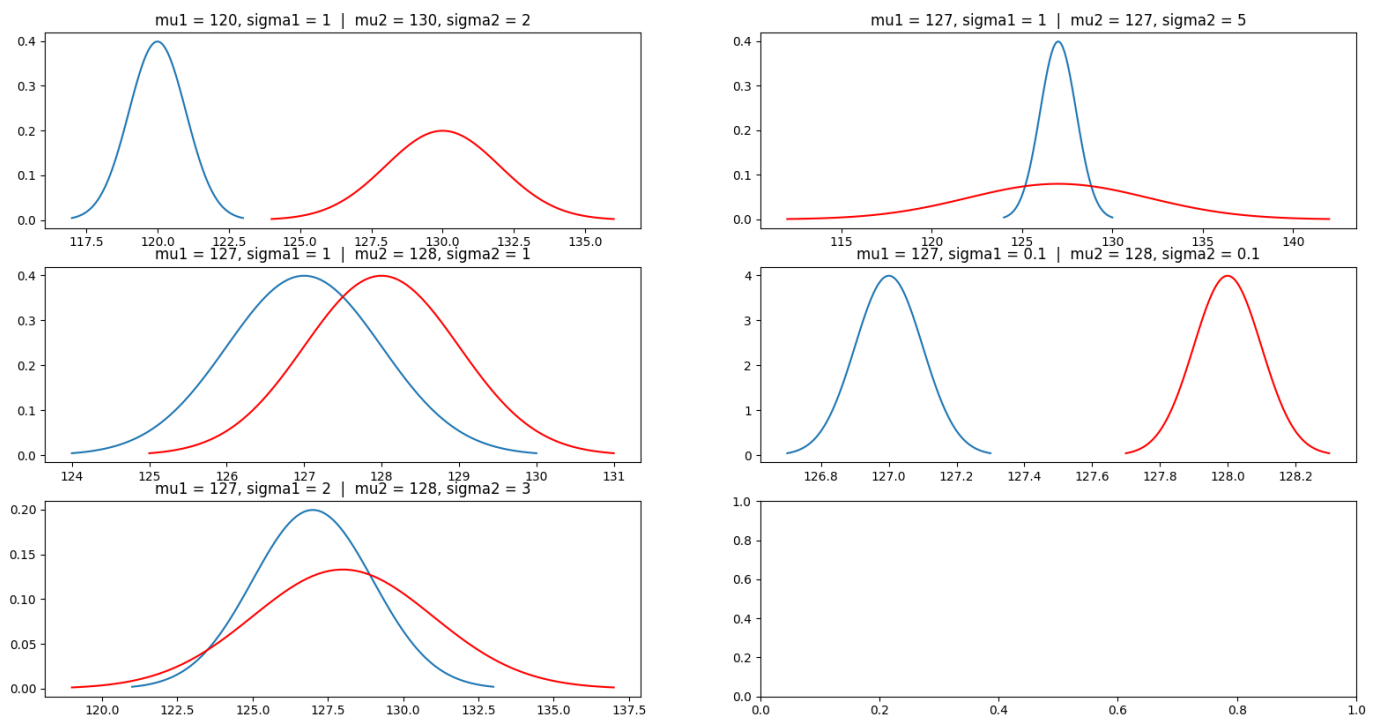
4. ET 5. TAUX POUR LES 6 SIGNAUX DIFFERENTS EN FONCTION DES PARAMETRES

Pour analyser les niveaux de bruit nous avons utilisé 6 signaux de tailles différentes et nous avons simulé du bruit avec 5 paramètres différents pour les mus et les sigmas. Voici un graph représentant toutes les gaussiennes pour simuler le bruit :

En bleu : μ_1 et σ_1

En rouge : μ_2 et σ_2

Gaussian distributions



Après la simulation nous avons fait un tableau représentant le taux d'erreurs moyen pour chaque signaux et chaque paramètres :

	s1: 50	s2: 250	s3: 50000	s4: 50000	s5: 50000	s6: 50000
m1: 120, s1: 1, m2: 130, s2: 2	0.0	0.0004	0.0005	0.0004	0.0004	0.0004
m1: 127, s1: 1, m2: 127, s2: 5	0.199	0.193	0.1763	0.1783	0.1773	0.1749
m1: 127, s1: 1, m2: 128, s2: 1	0.3332	0.2916	0.3085	0.3089	0.3101	0.3097
m1: 127, s1: 0.1, m2: 128, s2: 0.1	0.0	0.0	0.0	0.0	0.0	0.0
m1: 127, s1: 2, m2: 128, s2: 3	0.399	0.3754	0.3806	0.3814	0.3797	0.3789

Après la lecture de ce tableau, nous pouvons remarquer que le taux d'erreur est plus fort lorsque les gaussiennes simulant le niveau de bruit se chevauchent. Plus le taux d'erreur est grand plus le bruit est fort.

Un fort niveau de bruit peut correspondre à un taux d'erreur à 10% (soit 0.1) et un faible niveau de bruit peut correspondre à un taux d'erreur inférieure à 10%.

2 APPORT DES METHODES BAYESIENNES DE SEGMENTATION

2. TAUX POUR LES 6 SIGNAUX DIFFERENTS EN FONCTION DES PARAMETRES ET COMPARAISON DEUX METHODES DE CALCULES.

Nous constatons que, plus la longueur du signal augmente, plus le delta entre le calcul de l'erreur par la méthode bayésienne et mpm diminue.

	s1: 50	s2: 250	s3: 50000	s4: 50000	s5: 50000	s6: 50000
m1: 120, s1: 1, m2: 130, s2: 2	0.0	0.0	0.0	0.0	0.0	0.0
m1: 127, s1: 1, m2: 127, s2: 5	0.0002	0.0011	0.0	0.0	0.0	0.0
m1: 127, s1: 1, m2: 128, s2: 1	0.054	0.0014	0.0001	0.0	0.0	0.0001
m1: 127, s1: 0.1, m2: 128, s2: 0.1	0.0	0.0	0.0	0.0	0.0	0.0
m1: 127, s1: 2, m2: 128, s2: 3	0.1117	0.004	0.0001	0.0001	0.0	0.0001

4. TAUX POUR LES 5 SIGNAUX SIMULES DIFFERENTS EN FONCTION DES PARAMETRES ET COMPARAISON DEUX METHODES DE CALCULES.

	s1: 500	s2: 500	s3: 500	s4: 500	s5: 500
m1: 120, s1: 1, m2: 130, s2: 2	0.0	0.0	0.0	0.0	0.0
m1: 127, s1: 1, m2: 127, s2: 5	0.061	0.0013	0.0035	0.0052	0.0354
m1: 127, s1: 1, m2: 128, s2: 1	0.1246	0.0191	0.0061	0.0422	0.1485
m1: 127, s1: 0.1, m2: 128, s2: 0.1	0.0	0.0	0.0	0.0	0.0
m1: 127, s1: 2, m2: 128, s2: 3	0.301	0.0695	0.0005	0.0277	0.1126

Pour 5 signaux simulées avec ($p_{11} = 0.2$, $p_{21} = 0.8$ / $p_{12} = 0.4$, $p_{22} = 0.6$ / $p_{13} = 0.5$, $p_{23} = 0.5$ / $p_{14} = 0.6$, $p_{24} = 0.4$ / $p_{15} = 0.8$, $p_{25} = 0.2$) nous constatons que le delta entre le calcul de l'erreur par la méthode bayésienne et mpm diminue quand, les signaux sont simulés avec des probabilités équivalentes