On cherche à estimer V (paramètre constant).
On relève la position du véhicule le long du rail à des instants $t_n = nT$.
On considère $Y_n = nTV + B_n$, avec B_n $N(0, \sigma_B^2)$.
À $t_n = t_0 = 0$, le mobile se trouve en 0.

TD1 Expériences

1. On trace la droite qui passe au mieux par tous les points et l'origine, on trouve une pente d'environ 1 m/s.

   L'hypothèse "bruit blanc" a l'air de marcher mais ne pas faire de conclusion rapide (on n'a que 10 mesures).

2. La "meilleure droite" ne passe pas par l'origine. Les causes possibles sont : un bruit non centré, ou une position non nulle à $t_0 = 0$. On a l'impression que le bruit est corrélé, mais on ne peut pas tirer de conclusion.

3. Pour obtenir la ddp de $\hat{V}$:
   - Méthode basée sur l'expérience : chaque jeu d'observation donne $\hat{v}_i$ et on trace l'histogramme (voir TP d'initiation à Matlab).
   - Méthode de changement de variable : ddp de B_n, puis ddp de Y_n et enfin (passage difficile) ddp de $\hat{V}$.

   Remarque : $\hat{V}\ N(m_{\hat{v}}, \sigma^2_{\hat{V}})$
   - 1er estimateur (non biaisé) : $m_{\hat{v}} = 1m/s$ et $\sigma_{\hat{V}} = 0,08m/s$
   - 2ème estimateur (non biaisé) : $m_{\hat{v}} = 1m/s$ et $\sigma_{\hat{V}} = 0,04m/s$, meilleur estimateur car meilleur écart-type.

TD1.1 Estimateur empirique

Dans cette partie, $Y_n = nTV + B_n$ avec bruit faible, donc $V = \frac{y_n}{nT}$.

1. On mesure y_n.

$$\hat{V}_{emp} = \frac{Y_n}{nT}$$

$$E[\hat{V}_{emp}] = \frac{E[Y_n]}{nT} = \frac{E[nTV + B_n]}{nT} = V + \frac{E[B_n]}{nT} = V$$

   donc l'estimateur est non biaisé.

$$\sigma^2_{\hat{V}_{emp}} = E[(\hat{V}_{emp} - m_{\hat{V}_{emp}})^2] = E[(\frac{Y_n}{nT} - V)^2] = E[(\frac{nTV + B_n}{nT} - V)^2] = \frac{1}{(nT)^2}E[B_n^2]$$

$$\sigma_{\hat{V}_{emp}} = \frac{\sigma_B}{nT}$$

   Ainsi, pour minimiser $\sigma_{\hat{V}_{emp}}$, on prend $n = N$ (la plus grande mesure).

2. On dispose de N mesures $y_1, ...y_N$

$$\hat{V}_{emp} = \frac{\sum_n \frac{Y_n}{nT}}{n}$$

L'estimateur n'est pas biaisé car :

$$E[\hat{V}_{emp}] = \frac{\sum_n \frac{E[Y_n]}{nT}}{N} = V$$

Écart-type de l'estimateur :

$$\begin{aligned}
\sigma^2_{\hat{V}_{emp}} &= E[(\hat{V}_{emp} - m_{\hat{V}_{emp}})^2] \\
&= E[(\hat{V}_{emp} - V)^2] \\
&= E[(\frac{\sum_n V + \frac{B_n}{nT}}{N} - \frac{NV}{N})^2] \\
&= \frac{1}{N^2} E[(\sum_n \frac{B_n}{nT})^2] \\
&= \frac{1}{(NT)^2} E[(\sum_n \frac{B_n}{n})^2] \\
&= \frac{1}{(NT)^2} (\sum_n \frac{E[B_n^2]}{n^2} + \sum_{n \neq m} \frac{E[B_n B_m]}{nm})
\end{aligned}$$

Le bruit est blanc, donc les $E[B_n B_m] = 0$

$$\begin{aligned}
\sigma^2_{\hat{V}_{emp}} &= \frac{\sigma_B^2}{(NT)^2} \sum_n \frac{1}{n^2} \\
\sigma_{\hat{V}_{emp}} &= \frac{\sigma_B}{nT} \sqrt{S_1}
\end{aligned}$$

Or, $\frac{\sigma_B}{nT}\sqrt{S_1} > \frac{\sigma_B}{nT}$. Cela signifie que notre estimateur avec 1 mesure est "meilleur" que celui avec n mesures. On est triste d'avoir considéré les premières mesures qui sont très sensibles, comme nous, mais au bruit.

TD1.2 Préliminaires

V grandeur certaine mais inconnue, $Y_n = nTV + B_n$ et $B_n = N(0, \sigma_B^2)$.

1. Moyenne de Y_n : $m_n = E[Y_n] = nTV$.
   Écart-type de Y_n : $\sigma_n^2 = E[(Y_n - m_n)^2] = E[B_n^2]$ donc $\sigma_n = \sigma_B$
   Coefficient de corrélation :

$$\rho_{mn} = \frac{E[(Y_n - m_n)(Y_m - m_m)]}{\sigma_n \sigma_m} = \frac{E[B_n B_m]}{\sigma_B^2} = \begin{cases} 1 & \text{si} \quad m = n \\ 0 & \text{si} \quad m \neq n \end{cases} = \delta_{n-m}$$

2. Par changement de variables aléatoires si ça t'amuse,

$$f_{Y_n}(y_n) = f_{B_n}(y_n - nTV) = frac1\sqrt{2\pi\sigma_b^2} e^{-frac(y_n - nTV)^2 2\sigma_b^2}$$

Le caractère gaussien se conserve par transformation linéaire, i.e. toute combinaison linéaire de VA suivant une loi gaussienne suit aussi une loi gaussienne. Attention, ne pas sommer les ddp.

Les Y_i étant indépendants (car les B_i sont indépendants car blancs) :

$$f_{Y_1,...Y_n}(y_1,...y_n) = \prod_{i=1}^{N} f_{Y_i}(y_i) = \frac{1}{(2\pi)^{N/2}\sigma_B^N} \exp(-\frac{1}{2}\frac{\sum_n (y_n - nTV)^2}{\sigma_B^2})$$

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{(2\pi)^{N/2}\sigma_B^N} \exp(-\frac{1}{2}\frac{(\mathbf{y} - E[\mathbf{y}])^T(\mathbf{y} - E[\mathbf{y}])}{\sigma_B^2})$$

Rappel : la décorrélation n'implique pas l'indépendance, il faut le caractère "gaussiens dans leur ensemble".

TD1.3 Estimateur des moindres carrés

1. L'estimateur au sens des moindres carrés de V est :

$$\hat{v}_{MC} = \arg_V \min \sum_{n=1}^{N} (y_n - nTV)^2$$

En posant $J_{MC} : V \to \sum_{n=1}^{N} (y_n - nTV)^2$, J_{MC} est une parabole (concavité tournée vers le haut). On a alors une condition nécessaire et suffisante à la minimisation :

$$\begin{aligned}
\frac{dJ_{MC}}{dV}|_{V=\hat{V}_{MC}} = 0 &\Leftrightarrow \sum_{n=1}^{N} 2(-nT)(y_n - nTV) = 0 \\
&\Leftrightarrow \sum_{n=1}^{N} ny_n - \hat{v}_{MC}T \sum_{n=1}^{N} n^2 = 0 \\
&\Leftrightarrow \hat{v}_{MC} = \frac{\sum_n ny_n}{T\sum_{n=1}^{N} n^2} \\
&\Leftrightarrow \hat{V}_{MC} = \frac{\sum_n nY_n}{TS_2}
\end{aligned}$$

2. Calculons la moyenne de l'estimateur.

$$m_{MC} = E[\hat{v}_{MC}] = \frac{\sum_n nE[Y_n]}{TS_2} = \frac{\sum_n n(nTV)}{TS_2} = V$$

L'estimateur non biaisé.

On s'intéresse à son écart-type.

$$\begin{aligned}
\sigma_{MC}^2 &= E[(\hat{V}_{MC} - m_{MC})^2] \\
&= E[(\hat{V}_{MC} - V)^2] \\
&= E[(\frac{\sum_n nY_n}{TS_2} - V)^2] \\
&= E[(\frac{\sum_n n^2(TV + nB_n) - TV\sum_n n^2}{TS_2})^2] \\
&= E[(\frac{\sum_n nB_n}{TS_2})^2]
\end{aligned}$$

Comme les B_n sont décorrélés, les doubles produits sont tous nuls

$$\sigma_{MC}^2 = \frac{\sum_n n^2 E[B_n^2]}{T^2 S_2^2}$$
$$\sigma_{MC} = \frac{\sigma_B}{T\sqrt{S_2}}$$

Comme $S_2 = \sum_n n^2 > N^2$, on a $\sqrt{S_2} > N$ donc $\sigma_{MC} < \sigma_{emp} = \frac{\sigma_B}{NT}$.
Notre estimateur est meilleur que l'estimateur empirique.

TD1.4 Estimateur du maximum de vraisemblance

1. L'estimateur du maximum de vraisemblance de V est donné par

$$\hat{v}_{MV} = \arg_V \max f_{\mathbf{Y}}(\mathbf{y})$$

Remarque : dans le cas où V est incertain, $\arg_v \max f_{\mathbf{Y}/V=v}(\mathbf{y})$.

Y suit une loi gaussienne donc :

$$\hat{v}_{MV} = \arg_V \max f_{\mathbf{Y}}(\mathbf{y}) = \arg_V \min \sum_n (y_n - nTV)^2 = \arg\min J_{MC}(V)$$

2. Identique à la partie précédente car $\hat{V}_{MV} = \hat{V}_{MC}$.

TD1.5 Estimateur du maximum a posteriori

- V suit une loi gaussienne $N(V_0, \sigma_V^2)$
- les VA B_n et V sont indépendantes 2 à 2

On a 2 types d'informations :

- celle qui vient des mesures y_n
- celle qui vient de l'a priori V

1. Dans le cas où $V = v$ (v est certain), on ne change pas pour autant le comportement de $B_1, ...B_n$, donc de $Y_1...Y_n$:

$$f_{\mathbf{Y}/V=v}(\mathbf{y}) = \frac{1}{(2\pi)^{N/2}\sigma_B^N} \exp(-\frac{1}{2}\frac{\sum_n (y_n - nTv)^2}{\sigma_B^2})$$

2. On utilise la règle de Bayes :

$$f_{V/\mathbf{Y}=\mathbf{y}}(v) = \frac{f_{\mathbf{Y}/V=v}(\mathbf{y}) f_V(v)}{f_{\mathbf{Y}}(\mathbf{y})}$$

$f_{\mathbf{Y}}(\mathbf{y})$ ne dépend pas de v car on peut la calculer selon $f_{\mathbf{Y}}(\mathbf{y}) = \int_{\mathbb{R}} f_{\mathbf{Y},V}(\mathbf{y}, v) dv$.

3. L'estimateur du maximum a posteriori de V est donné par :

$$\hat{v}_{MAP} = \arg_v \max f_{V/\mathbf{Y}=\mathbf{y}}(v)$$

Or

$$f_{V/\mathbf{Y}=\mathbf{y}}(v) = cste \times \exp(-\frac{1}{2}(\frac{\sum_n (y_n - nvT)^2}{\sigma_B^2} + \frac{(v - V_0)^2}{\sigma_V^2}))$$

On pose $J_{MAP} = \frac{\sum_n (y_n - nvT)^2}{\sigma_B^2} + \frac{(v-V_0)^2}{\sigma_V^2}$ et on a alors $\hat{v}_{MAP} = \arg_v \min J_{MAP}(v)$.

CNS de maximisation :

$$\frac{dJ_{MAp}}{dV}|_{V=\hat{v}_{MAP}} = 0 \Leftrightarrow -2T\frac{\sum_n(y_n - nvT)n}{\sigma_B^2} + 2\frac{(v-V_0)}{\sigma_V^2} = 0$$

$$\Leftrightarrow \hat{v}_{MAP} = \frac{\frac{T\sum_n ny_n}{\sigma_B^2} + \frac{V_0}{\sigma_V^2}}{\frac{T^2\sum_n n^2}{\sigma_B^2} + \frac{1}{\sigma_V^2}}$$

$$\Leftrightarrow \hat{v}_{MAP} = \frac{\frac{\hat{v}_{MV}}{\sigma_{MV}^2} + \frac{V_0}{\sigma_V^2}}{\frac{1}{\sigma_{MV}^2} + \frac{1}{\sigma_V^2}}$$

Donc

$$\hat{V}_{MAP} = \frac{\frac{\hat{V}_{MV}}{\sigma_{MV}^2} + \frac{V_0}{\sigma_V^2}}{\frac{1}{\sigma_{MV}^2} + \frac{1}{\sigma_V^2}}$$

C'est un barycentre entre les mesures représentées par $\hat{v}_{MV}$ et l'a priori V_0.

- Si σ_V "petit", alors $\hat{V}_{MAP} \approx V_0$: beaucoup d'a priori donc on n'a pas exploité les mesures.
- Si σ_V "grand", alors $\hat{V}_{MAP} \approx \hat{V}_{MV}$: l'a priori est tellement pourri qu'on n'en tient pas compte.

4. On forme l'erreur d'estimation

$$\tilde{V}_{MAP} = \hat{V}_{MAP} - V$$

- Biais ?

$$E[\tilde{V}_{MAP}] = E[\hat{V}_{MAP}] - E[V] = 0$$

- Variance de l'erreur d'estimation : puissance de l'erreur dans le cas non biaisé.

$$\begin{aligned}
\sigma_{MAP}^2 &= E[(\tilde{V}_{MAP} - E[\hat{V}_{MAP}])^2] \\
&= E[(\hat{V}_{MAP} - V)^2] \\
&= E[(\frac{\frac{\hat{V}_{MV}-V}{\sigma_{MV}^2} + \frac{V_0-V}{\sigma_V^2}}{\frac{1}{\sigma_{MV}^2} + \frac{1}{\sigma_V^2}})^2] \\
&= \frac{1}{(\frac{1}{\sigma_{MV}^2} + \frac{1}{\sigma_V^2})^2} E[(\frac{\hat{V}_{MV}-V}{\sigma_{MV}^2} + \frac{V_0-V}{\sigma_V^2})^2] \\
&= \frac{1}{(\frac{1}{\sigma_{MV}^2} + \frac{1}{\sigma_V^2})^2} (\frac{E[(\hat{V}_{MV}-V)^2]}{\sigma_{MV}^4} + \frac{E[(V_0-V)^2}{\sigma_V^4}) \\
&= \frac{1}{(\frac{1}{\sigma_{MV}^2} + \frac{1}{\sigma_V^2})^2} (\frac{\sigma_{MV}^2}{\sigma_{MV}^4} + \frac{\sigma_V^2}{\sigma_V^4}) \\
\sigma_{MAP}^2 &= (\frac{1}{\sigma_{MV}^2} + \frac{1}{\sigma_V^2})^{-1}
\end{aligned}$$

On a donc $\sigma_{emp} > \sigma_{MV} > \sigma_{MAP}$.