Logique de décision de validation de données

November 3, 2016

Dans les expressions algébriques données dans ce document :

- le symbole **point** (**.**) représente l'opérateur "**produit scalaire**",
- le symbole **croix** ($\times$) représente l'opérateur "**produit matriciel**".

Introduction

Pour la validation des données, nous devons mettre en place un algorithme permettant de discriminer les données valides des données invalides. Chaque composant validera ou invalidera les données des autres calculateurs, en considérant sa donnée comme valeur "valide" de référence et en considérant un écart tolérable ε. La validation de chaque composant i sera donnée dans la ligne i de la matrice de validation V. En fonction de la confiance du composant i envers chaque autre composant et en fonction d'un égo aléatoire γ (une réalisation par échantillonnage), un consensus sera calculé et déterminera l'avis retenu par le composant i, qui lui permettra de mettre à jour son taux de confiance envers les autres composants (cet avis sera propre au composant i et ne sera pas communiqué aux autres composants, y compris l'arbitre : il sert uniquement à faire évoluer le taux de confiance propre au composant i).

1 Description de l'algorithme

Déterminer une tolérance d'erreur admissible. Par exemple, pour une commande sur un moteur dont l'amplitude est sur l'intervalle [0;10] V, on peut considérer qu'une erreur de 1% FS (Full Scale = Pleine Echelle) soit $10 \times 1\% = 0.1$ est admissible.

1. La donnée de référence considérée comme valide est celle du composant qui effectue la validation.
2. On discrimine les données par rapport à leur écart vis-à-vis de cette donnée de "référence" : si l'écart est inférieur (ou égal) au seuil de tolérance fixé, la donnée est considérée valide, sinon elle est considérée invalide. Une donnée absente est également considérée invalide.
3. On détermine un consensus en pondérant la validation de chaque composant par le taux de confiance du composant i envers les composants dont est issu cette validation. Un égo aléatoire (multiplicateur du taux de confiance du composant envers lui-même) pour lui donner la possibilité de donner parfois un avis "différent" de l'avis général, ce qui permettra une diversification des taux de confiance en cas d'ambiguïté. Cet égo γ_i sera tel que $\gamma_i \in [1; 1.5]$ et sera donné par une nouvelle réalisation à chaque échantillonnage.
4. Le consensus permettra par la suite de modifier le taux de confiance propre au composant i calculant la validation.

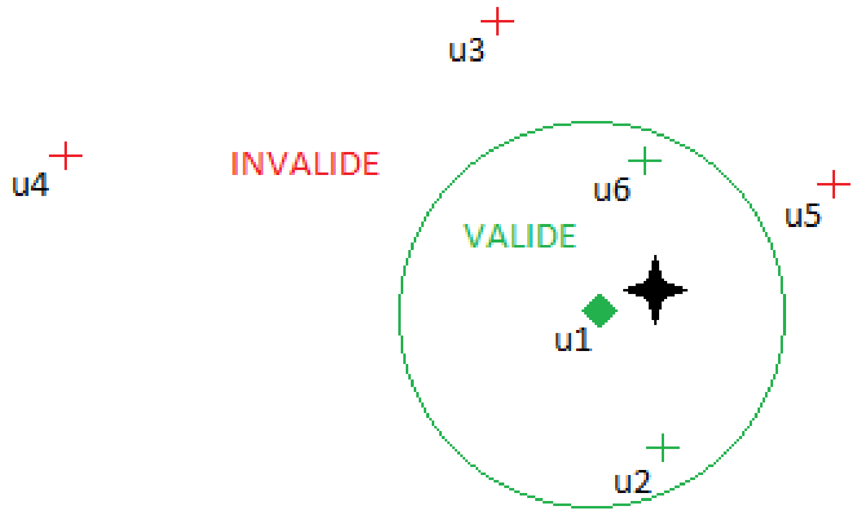

Figure 1: Validation de données

Formule :

$$\boxed{\hat{v}_i = \big((\Gamma_i \times M)\,.\,(\gamma_i.\Gamma_i + \overline{\Gamma_i})\big) \times V}$$

On construit v_i, de coefficients $e_{i,j}$, à partir des coefficients $\hat{e}_{i,j}$ de $\hat{v}_i$ avec :

$$\begin{cases} e_{i,j} = 1, & \hat{e}_{i,j} \geq 0 \\ e_{i,j} = -1, & \hat{e}_{i,j} < 0 \end{cases}$$

2 Code Matlab

Le code Matlab pour implémenter cette fonction de consensus est la suivante :

```
function [v] = validation_consensus(M,V,i,ego)
% Paramètres :
% M = matrice de confiance,
% V = matrice de validation,
% i = indice du composant effectuant le consensus de validation
% ego = coefficient multiplicateur du taux de confiance du composant i
%
% Sortie :
% v = vecteur de validation {1 = donnée valide, -1 = donnée invalide}

G=zeros(1,6);       G(i)=1;
Gc=ones(1,6);       Gc(i)=0;

v=(((G*M).*(ego*G+Gc))*V>=0).*2-1;

end
```

3 Exemple

- Considérons V une matrice de validation de chaque composant extraite de la Figure 1ci-dessus (remarque : V est toujours symétrique de par la méthode de validation) :

$$V = \begin{bmatrix} 1 & 1 & -1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 1 & -1 & -1 & 1 \\ -1 & -1 & -1 & 1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 \end{bmatrix}$$

- Considérons M une matrice de confiance des composants :

$$M = \begin{bmatrix} 1.0 & 1.0 & 0.9 & 1.0 & 0.2 & 0.1 \\ 1.0 & 1.0 & 1.0 & 1.0 & 0.2 & 0.1 \\ 0.9 & 0.9 & 1.0 & 1.0 & 0.2 & 0.2 \\ 1.0 & 0.9 & 0.8 & 0.9 & 0.3 & 0.2 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.8 & 0.7 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.8 & 0.9 \end{bmatrix}$$

Calculons la validation du composant 1 avec un égo aléatoire tiré $\gamma_1 = 1.5$:

$$\hat{v}_1 = \left((\Gamma_1 \times M) \, . \, (\gamma_1 . \Gamma_1 + \overline{\Gamma_1}) \right) \times V$$

$$\hat{v}_1 = \left(\begin{bmatrix} 1.0 & 1.0 & 0.9 & 1.0 & 0.2 & 0.1 \end{bmatrix} . \begin{bmatrix} 1.5 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \right) \times \begin{bmatrix} 1 & 1 & -1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 1 & -1 & -1 & 1 \\ -1 & -1 & -1 & 1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 \end{bmatrix}$$

$$\hat{v}_1 = \begin{bmatrix} 1.5 & 1.0 & 0.9 & 1.0 & 0.2 & 0.1 \end{bmatrix} \times \begin{bmatrix} 1 & 1 & -1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 1 & -1 & -1 & 1 \\ -1 & -1 & -1 & 1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 \end{bmatrix}$$

$$\hat{v}_i = \begin{bmatrix} 0.5 & 0.3 & -2.7 & -2.7 & -4.1 & 0.7 \end{bmatrix}$$

$$\Longrightarrow v_1 = \begin{bmatrix} 1 & 1 & -1 & -1 & -1 & 1 \end{bmatrix}$$

En appliquant le même raisonnement pour chaque composant, avec un taux d'égo constant de 2 (pour simplifier) on obtient la matrice de validation après consensus V_G (taux d'égo constant de 1.5) issu de V et M :

$$V_G = \begin{bmatrix} 1 & 1 & -1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ -1 & -1 & \mathbf{-1} & -1 & -1 & 1 \\ -1 & -1 & -1 & \mathbf{-1} & -1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 \end{bmatrix}$$

Pour rappel :

$$V = \begin{bmatrix} 1 & 1 & -1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 1 & -1 & -1 & 1 \\ -1 & -1 & -1 & 1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 \end{bmatrix}$$

En comparant à la matrice V initiale, on constate que 3 valeurs ont changées :

- la donnée du calculateur 3 après consensus est invalidée : en effet, le composant 3 fait confiance aux composants 1,2 et 4 qui déclarent la valeur donnée par le composant 3 (lui-même donc) invalide : malgré son égo, il tient compte de l'avis des autres et invalide la donnée,
- le même raisonnement s'applique pour la donnée du composant 4.
- la donnée du composant 6 était invalidée par le composant 4 alors qu'elle est validée après consensus. Le composant 4 indique, avec le composant 2 (en qui il a confiance), que la donnée du composant 6 est invalide. Mais les composants 1 et 3 (en qui il le composant 4 fait aussi confiance) l'invalident. C'est aussi le cas des composants 5 et 6, en qui le composant 4 à une faible confiance (mais une confiance non négligeable ici). Le composant 4 va donc pencher, malgré son égo, vers l'avis des composants 1,3,5 et 6 et valider la donnée du composant 6.

Poursuivons l'analyse pour la donnée du composant 2 : si l'égo est de 1.3 la donnée est validée pas les composants 1 et 2 (matrice suivante $M_{1.3}$). Si l'égo vaut 1.2, la donnée est invalidée par ces mêmes composants (matrice suivante $M_{1.2}$) :

$$M_{1.3} = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{-1} & \mathbf{-1} & \mathbf{-1} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & \mathbf{-1} & \mathbf{-1} & \mathbf{-1} & \mathbf{1} \\ -1 & -1 & -1 & -1 & -1 & 1 \\ -1 & -1 & -1 & -1 & -1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 \end{bmatrix}, \, M_{1,2} = \begin{bmatrix} \mathbf{1} & \mathbf{-1} & \mathbf{-1} & \mathbf{-1} & \mathbf{-1} & \mathbf{1} \\ \mathbf{1} & \mathbf{-1} & \mathbf{-1} & \mathbf{-1} & \mathbf{-1} & \mathbf{1} \\ -1 & -1 & -1 & -1 & -1 & 1 \\ -1 & -1 & -1 & -1 & -1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 \end{bmatrix}$$

D'après le cas étudié illustré Figure 1 dont on a déduit la matrice V et avec la matrice M considérée, cela paraît cohérent : la validité du composant 2 semble ambigüe et selon l'égo du composant 1 ou 2 on pourrait aussi bien considérer la valeur valide ou non !

Ici nous avons, encore une fois, considéré un cas extrême ou tout les composants renvoient des valeurs relativement éloignées et avec 2 composants potentiellement défaillants mais avec une forte confiance mutuelle. En règle générale, les composants ne devraient pas devenir défaillants simultanément, il est encore plus improbable qu'ils renvoient tous des valeurs différentes aussi éloignées et enfin, encore moins probable que la confiance de la population se scinde en 2 clans comme c'est le cas ici avec les composants $\{1, 2, 3, 4\}$ et $\{5, 6\}$: l'opinion d'un composant suivra presque systématiquement toujours la majorité (au moins avec un nombre de composants "sains" aussi important) car dès qu'un composant devient défaillant, l'ensemble des autres composants devraient l'exclure et son taux de confiance diminuera rapidement (malgré l'égo de chaque composant), ce qui diminuera d'autant plus son influence sur les avis futurs...