Surveillance Distribuée de Systèmes Multi-agents

UPJV, Département EEA

Master 2 3EA, EC53

UE alternants

Fabio MORBIDI

Laboratoire MIS
Équipe Perception Robotique
fabio.morbidi@u-picardie.fr

Mardi, Mercredi et Jeudi
13h30-16h00 (CM et TD salle CURI 305)
13h30-17h30 (TP salle CURI 305)

AU 2022-2023

UNIVERSITÉ de Picardie Jules Verne

Ch. 3: Estimation distribuée

- Introduction à la théorie des graphes **Partie 1**
- Protocole de consensus **Partie 2**
- Dynamic average consensus estimators **Partie 3**
- Méthode des moindres carrés distribuée **Partie 4**
- Filtre de Kalman distribué **Partie 5**

Méthode des moindres carrés distribuée

Introduction

- La méthode des moindres carrés *linéaires* peut être reformulée de *façon distribuée*

- Les estimateurs basés sur la méthode des moindres carrés ne font aucune hypothèse sur les propriétés statistiques du bruit qui entache la variable sous-jacente $\boldsymbol{\theta}$ (au début du Ch. 1, nous avons désigné ce vecteur de paramètres par le symbole $\boldsymbol{\beta}$)
  - Pour cette raison, ils sont facilement applicables à un large éventail de problèmes d'estimation distribuée

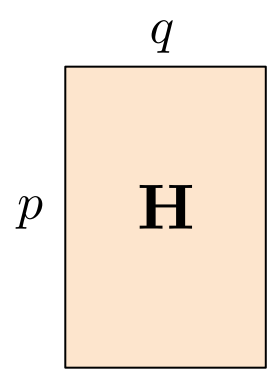
Méthode des moindres carrés

- Le modèle d'observation consiste en une *fonction linéaire* de la variable $\boldsymbol{\theta} \in \mathbb{R}^q$ qui est entachée par le bruit additif $\mathbf{r}$:

$$\mathbf{z} = \mathbf{H}\,\boldsymbol{\theta} + \mathbf{r}$$

avec $\mathbf{z}, \mathbf{r} \in \mathbb{R}^p$ et $\mathbf{H} \in \mathbb{R}^{p \times q}$ $(p > q)$.

  - Chaque composante z_i du vecteur $\mathbf{z}$ s'appelle canal de mesure
  - $\mathbf{H}$ est la matrice d'observation. On fera l'hypothèse que $\mathbf{H}$ soit une matrice de rang ligne égal à q. Cette condition de rang garantit que les canaux de mesure *ne sont pas entièrement redondants*



Méthode des moindres carrés

- En l'*absence d'informations* sur les statistiques du bruit de mesure $\mathbf{r}$, l'estimation par la méthode des moindres carrés procède à la minimisation de la fonction de coût suivante:

$$S(\boldsymbol{\theta}) = (\mathbf{z} - \mathbf{H}\boldsymbol{\theta})^T (\mathbf{z} - \mathbf{H}\boldsymbol{\theta})$$

- Comme S est une fonction de $\boldsymbol{\theta}$ *différentiable* et *convexe*, on peut calculer la valeur minimale en égalant son gradient à zéro, à savoir:

$$\frac{\partial\, S(\boldsymbol{\theta})}{\partial\, \boldsymbol{\theta}} = \mathbf{0}$$

- On trouve ainsi que:

$$\widehat{\boldsymbol{\theta}} = \underbrace{(\mathbf{H}^T \mathbf{H})^{-1}\, \mathbf{H}^T}_{\mathbf{H}^\dagger} \mathbf{z}$$

Estimation par moindres carrées linéaires de $\boldsymbol{\theta}$

Pseudo-inverse (à gauche) de $\mathbf{H}$

Méthode des moindres carrés pondérés

- Si $\mathbf{r}$ est un bruit blanc gaussien à moyenne zéro avec matrice de covariance définie positive $\boldsymbol{\Sigma}$, à savoir $\mathbf{r} \sim \mathcal{N}(\mathbf{0},\, \boldsymbol{\Sigma})$, la minimisation de la *fonction de coût pondérée*:

$$S(\boldsymbol{\theta}) \;=\; (\mathbf{z} - \mathbf{H}\boldsymbol{\theta})^T\, \boldsymbol{\Sigma}^{-1}(\mathbf{z} - \mathbf{H}\boldsymbol{\theta})$$

produit l'estimé optimal de $\boldsymbol{\theta}$:

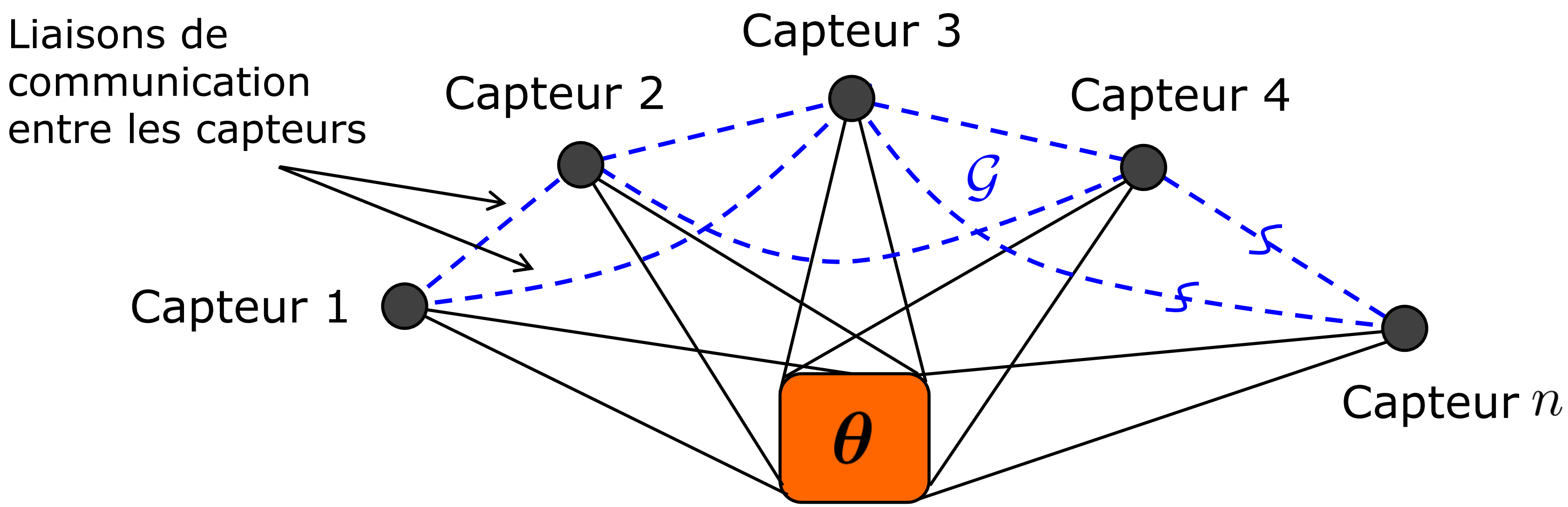
$$\widehat{\boldsymbol{\theta}} \;=\; \underbrace{(\mathbf{H}^T\, \boldsymbol{\Sigma}^{-1}\mathbf{H})^{-1}\, \mathbf{H}^T\, \boldsymbol{\Sigma}^{-1}}_{\mathbf{H}^{\dagger}}\mathbf{z}$$

Estimation par moindres carrés pondérés de $\boldsymbol{\theta}$

Pseudo-inverse pondérée (à gauche) de $\mathbf{H}$

- L'introduction de la matrice $\boldsymbol{\Sigma}^{-1}$ dans la fonction de coût $S(\boldsymbol{\theta})$ est motivée par le désir de *biaiser* l'estimé optimal vers ces mesures qui sont les moins incertaines

Moindres carrés dans un réseau de capteurs



- Supposons maintenant d'avoir n capteurs en réseau qui réalisent des mesures de $\boldsymbol{\theta}$. Le vecteur d'observation du capteur i est:

$$\mathbf{z}_i = \mathbf{H}_i \boldsymbol{\theta} + \mathbf{r}_i, \quad i \in \{1, \ldots, n\}$$

avec $\mathbf{z}_i \in \mathbb{R}^{p_i}$ et $\mathbf{H}_i \in \mathbb{R}^{p_i \times q}$.

- Par souci de simplicité, on fera l'hypothèse que les matrices de covariance des bruits de mesure $\mathbf{r}_i$ sont l'identité: $\boldsymbol{\Sigma}_i = \mathbf{I}_{p_i}$, $i \in \{1, \ldots, n\}$

Moindres carrés dans un réseau de capteurs

- Considérons maintenant la matrice d'observation *centralisée*:

$$\mathbf{H} = [\mathbf{H}_1^T \ \ \mathbf{H}_2^T \ \ \cdots \ \ \mathbf{H}_n^T]^T \in \mathbb{R}^{p \times q},$$

  avec $p = p_1 + p_2 + \ldots + p_n$

- En utilisant la matrice $\mathbf{H}$, nous pouvons écrire l'estimateur des moindres carrés *centralisé* comme suit:

$$\widehat{\boldsymbol{\theta}} = \left(\sum_{i=1}^{n} \mathbf{H}_i^T \mathbf{H}_i \right)^{-1} \left(\sum_{i=1}^{n} \mathbf{H}_i^T \mathbf{z}_i \right),$$

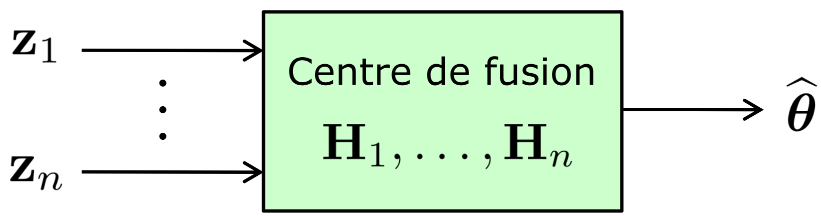
  pourvu que les bruits additifs $\mathbf{r}_i$ soient *statistiquement indépendants*

Moindres carrés dans un réseau de capteurs

- La **nature additive** de l'équation,

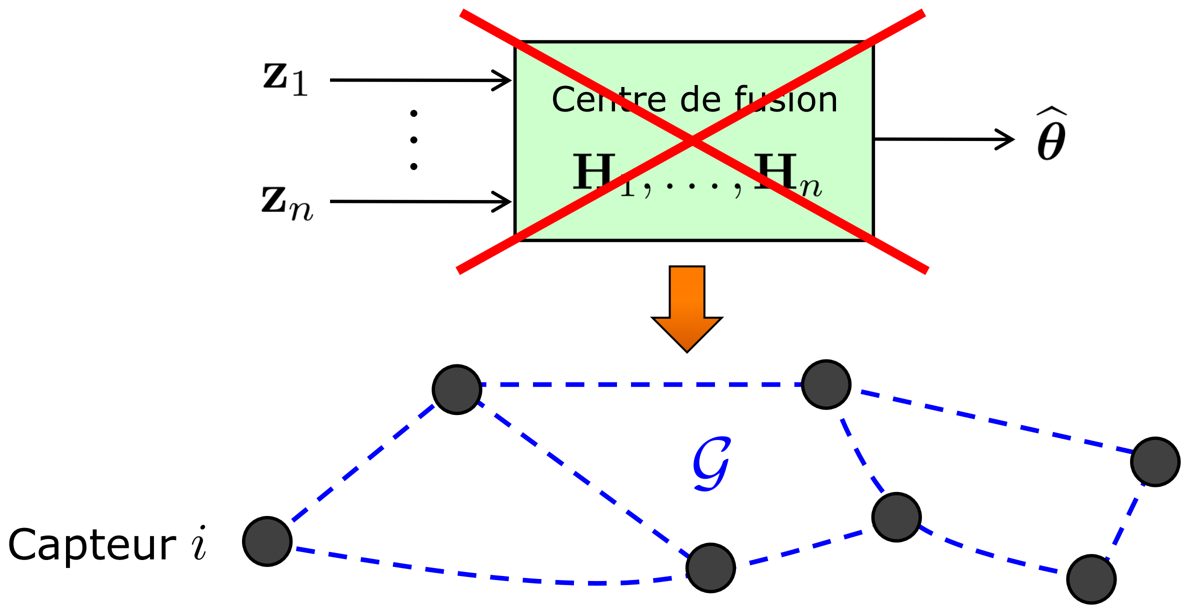
$$\widehat{\boldsymbol{\theta}} = \left(\sum_{i=1}^{n} \mathbf{H}_i^T \mathbf{H}_i\right)^{-1} \left(\sum_{i=1}^{n} \mathbf{H}_i^T \mathbf{z}_i\right),$$

suggère que si chaque capteur fournit la mesure brute $\mathbf{z}_i$ à un **centre de fusion** qui a une connaissance préalable de chaque matrice d'observation $\mathbf{H}_i$, alors le centre de fusion peut calculer efficacement l'estimé aux moindres carrés $\widehat{\boldsymbol{\theta}}$



Moindres carrés dans un réseau de capteurs

- Cependant, pour améliorer la *scalabilité* et la *modularité,* et pour avoir un estimateur *moins vulnérable aux pannes*, il est souhaitable de calculer $\widehat{\theta}$ **sans besoin** d'un centre de fusion
- Le **protocole de consensus** vient au secours, car il nous permet de concevoir un algorithme distribué pour les moindres carrés, qui utilise le réseau de communication $\mathcal{G}$ des capteurs pour calculer $\widehat{\theta}$ *sans besoin d'un centre de fusion*



Moindres carrés dans un réseau de capteurs

- Considérons le scénario *le plus simple possible*:
  - Estimation d'une variable ***scalaire*** $\theta \in \mathbb{R}$.
  - Les "matrices" d'observation sont $H_i = 1,\ \forall\, i$.
- Sous ces hypothèses, nous avons que l'équation de mesure est:

$$z_i = \theta + r_i, \;\; i \in \{1, \ldots, n\}$$

et

$$\widehat{\theta} = \frac{1}{n} \sum_{i=1}^{n} z_i$$

Estimation par moindres carrés centralisée de θ

- Cette dernière expression (une moyenne arithmétique) nous permet d'utiliser le *protocole de consensus* (en temps discret) pour calculer la solution du problème des moindres carrés de *façon distribuée*

Moindres carrés dans un réseau de capteurs

- Soit

$$\mathbf{W} = \mathrm{diag}(w_1, \ldots, w_m),$$

  une *matrice diagonale de pondération* où $w_j > 0$ est le poids de l'arête e_j du graphe $\mathcal{G}$ (indexé de manière cohérente, selon l'ordre des colonnes de la matrice d'incidence $\mathbf{D}(\mathcal{G})$ correspondante)

- Considérons maintenant l'itération suivante réalisée par le capteur i :

$$\widehat{\theta}_i(k+1) = \widehat{\theta}_i(k) + \Delta \sum_{j \in \mathcal{N}(i)} w_{ij}(\widehat{\theta}_j(k) - \widehat{\theta}_i(k)),\ i \in \{1, \ldots, n\},$$

  avec

  - $\widehat{\theta}_i(k)$: estimé local du capteur i de la variable θ à l'instant k
  - $\Delta \in (0,\, 1)$: incrément pour la mise à jour de l'estimé de θ
  - w_{ij} : poids de l'arête reliant le sommet i au sommet j
  - $\mathcal{N}(i)$: ensemble des sommets adjacents à i dans le graphe $\mathcal{G}$

Moindres carrés dans un réseau de capteurs

- Soit
$$\mathbf{L}_w(\mathcal{G}) \;=\; \mathbf{D}(\mathcal{G})\mathbf{W}\,\mathbf{D}^T(\mathcal{G}),$$
la *matrice laplacienne pondérée* de $\mathcal{G}$ et définissons la matrice des "itérations" (ou de Perron) suivante:
$$\mathbf{M}_w(\mathcal{G}) \;=\; \mathbf{I}_n - \Delta\,\mathbf{L}_w(\mathcal{G}).$$
- Nous pouvons ainsi récrire de façon compacte les itérations vues dans la diapo précédente comme:
$$\widehat{\boldsymbol{\theta}}(k+1) \;=\; \mathbf{M}_w(\mathcal{G})\,\widehat{\boldsymbol{\theta}}(k),$$
avec $\widehat{\boldsymbol{\theta}} \triangleq [\widehat{\theta}_1, \ldots, \widehat{\theta}_n]^T$.
- Donc
$$\widehat{\boldsymbol{\theta}}(k) \;=\; \mathbf{M}_w^k(\mathcal{G})\,\widehat{\boldsymbol{\theta}}(0), \quad k \in \{1, 2, \ldots\},$$
où $\widehat{\boldsymbol{\theta}}(0)$ est l'estimé initial de $\boldsymbol{\theta}$.
- La convergence de cette séquence dépend du comportement des puissances de $\mathbf{M}_w(\mathcal{G})$, qui dépend à son tour du spectre de $\mathbf{M}_w(\mathcal{G})$

Moindres carrés dans un réseau de capteurs

Proposition (*Résultat principal*)
Soit la séquence,

$$\widehat{\boldsymbol{\theta}}(k) \;=\; \mathbf{M}_w^k(\mathcal{G})\,\widehat{\boldsymbol{\theta}}(0), \quad k \in \{1, 2, \ldots\},$$

initialisée avec $\widehat{\boldsymbol{\theta}}(0) \;=\; [z_1,\, z_2,\, \ldots,\, z_n]^T$, où z_i est l'observation du capteur i. Alors,

$$\lim_{k\to\infty} \widehat{\boldsymbol{\theta}}(k) \;=\; \left(\frac{1}{n}\sum_{i=1}^{n} z_i\right)\mathbf{1},$$

si et seulement si le réseau $\mathcal{G}$ est *connexe* et la valeur propre *la plus forte* de $\mathbf{L}_w(\mathcal{G})$ satisfait la condition:

$$\lambda_n(\mathbf{L}_w(\mathcal{G})) < \frac{2}{\Delta}$$

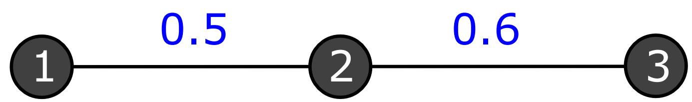
<u>*Preuve*</u> : Elle ressort de l'observation que le spectre de $\mathbf{M}_w(\mathcal{G})$ est simplement:

$$1 - \Delta\,\lambda_i(\mathbf{L}_w(\mathcal{G})),\;\; i \in \{1, \ldots, n\}$$

Moindres carrés dans un réseau de capteurs

Exercice:

- Soit le graphe chaîne pondéré suivant ($\mathcal{G} = P_3$):



- Vérifier que la matrice laplacienne pondérée de $\mathcal{G}$ et la matrice des itérations correspondante, pour $\Delta = 1$, sont:

$$\mathbf{L}_w(\mathcal{G}) = \begin{bmatrix} 0.5 & -0.5 & 0 \\ -0.5 & 1.1 & -0.6 \\ 0 & -0.6 & 0.6 \end{bmatrix}, \quad \mathbf{M}_w(\mathcal{G}) = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.5 & -0.1 & 0.6 \\ 0 & 0.6 & 0.4 \end{bmatrix}.$$

- Vérifier que la valeur propre la plus forte de $\mathbf{L}_w(\mathcal{G})$ satisfait la condition:

$$\lambda_3(\mathbf{L}_w(\mathcal{G})) < \frac{2}{\Delta} = 2$$

  et que chaque élément de la matrice $\mathbf{M}_w^k(\mathcal{G})$ converge vers 1/3 lorsque $k \to \infty$

Moindres carrés dans un réseau de capteurs

Corollaire (*Choix de la matrice de pondération* $\mathbf{W}$)
Supposons que le j-ème élément diagonal de la matrice $\mathbf{W}$, qui correspond au poids sur l'arête $e_j = uv$, soit défini comme suit:

$$[\mathbf{W}]_{jj} = \frac{1}{\max\{d(u),\, d(v)\}},$$

où $d(u)$ dénote le degré du sommet u. Alors, pour tout $\Delta \in (0,\, 1)$, l'itération,

$$\widehat{\theta}_i(k+1) = \widehat{\theta}_i(k) + \Delta \sum_{j \in \mathcal{N}(i)} w_{ij}(\widehat{\theta}_j(k) - \widehat{\theta}_i(k)),$$

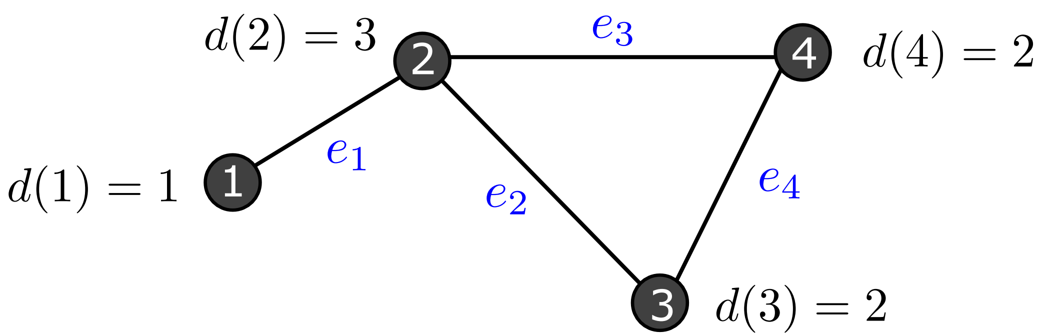
converge vers la valeur calculée par l'estimateur des moindres carrés centralisé

Remarque:
Dans la littérature, la formule précédente est connue sous le nom de *règle de ponderation de Metropolis–Hastings*

TD3, ex. IV.1
Fichier Matlab:
MoinCar_Distr.m

Moindres carrés dans un réseau de capteurs

<u>Exemple</u> (Règle de Metropolis-Hastings):



- Les poids sur les quatre arêtes e_1, e_2, e_3, e_4 sont:

$$
\begin{aligned}
[\mathbf{W}]_{11} &= \frac{1}{\max\{1,3\}} = \frac{1}{3} \\
[\mathbf{W}]_{22} &= \frac{1}{\max\{3,2\}} = \frac{1}{3} \\
[\mathbf{W}]_{33} &= \frac{1}{\max\{3,2\}} = \frac{1}{3} \\
[\mathbf{W}]_{44} &= \frac{1}{\max\{2,2\}} = \frac{1}{2}
\end{aligned}
$$

- La matrice de pondération cherchée est donc:

$$\mathbf{W} = \mathrm{diag}\left(1/3,\ 1/3,\ 1/3,\ 1/2\right)$$

Moindres carrés dans un réseau de capteurs : cas vectoriel

- On peut généraliser le résultat précédent pour une **variable vectorielle** $\boldsymbol{\theta} \in \mathbb{R}^q$.
- On fera l'hypothèse que la matrice d'observation $\mathbf{H}_i \in \mathbb{R}^{p_i \times q}$ du capteur i soit arbitraire, pourvu que $\mathbf{H} = [\mathbf{H}_1^T \;\; \mathbf{H}_2^T \;\; \cdots \;\; \mathbf{H}_n^T]^T$ soit de *rang plein ligne*
- Chaque capteur i stocke deux quantités, la matrice $\mathbf{P}_i \in \mathbb{R}^{q \times q}$ et le vecteur $\widehat{\boldsymbol{\theta}}_i \in \mathbb{R}^q$, et il effectue *localement* les itérations suivantes:

$$\left.\begin{aligned} \mathbf{P}_i(k+1) &= \mathbf{P}_i(k) + \Delta \sum_{j \in \mathcal{N}(i)} w_{ij}(\mathbf{P}_j(k) - \mathbf{P}_i(k)), \\ \widehat{\boldsymbol{\theta}}_i(k+1) &= \widehat{\boldsymbol{\theta}}_i(k) + \Delta \sum_{j \in \mathcal{N}(i)} w_{ij}(\widehat{\boldsymbol{\theta}}_j(k) - \widehat{\boldsymbol{\theta}}_i(k)), \end{aligned}\right. \quad i \in \{1, \ldots, n\},$$

avec l'initialisation,

$$\mathbf{P}_i(0) = \mathbf{H}_i^T \mathbf{H}_i, \quad \widehat{\boldsymbol{\theta}}_i(0) = \mathbf{H}_i^T \mathbf{z}_i, \quad i \in \{1, \ldots, n\},$$

où $\mathbf{z}_i \in \mathbb{R}^{p_i}$ dénote le vecteur d'observation du capteur i

Moindres carrés dans un réseau de capteurs : cas vectoriel

- Nous avons que (cf. la discussion dans le cas scalaire):

$$\lim_{k\to\infty} \mathbf{P}_i(k) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{H}_i^T \mathbf{H}_i,$$

$$\lim_{k\to\infty} \widehat{\boldsymbol{\theta}}_i(k) = \frac{1}{n}\sum_{i=1}^{n} \widehat{\boldsymbol{\theta}}_i(0) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{H}_i^T \mathbf{z}_i,$$

si l'incrément Δ et les poids w_{ij} respectent les conditions imposées par la Proposition vue dans le cas scalaire

- Chaque capteur peut ainsi calculer *localement* l'estimé fourni par la méthode des moindres carrés centralisée. En fait, nous avons que:

$$\widehat{\boldsymbol{\theta}} = \lim_{k\to\infty} \mathbf{P}_i^{-1}(k)\, \widehat{\boldsymbol{\theta}}_i(k).$$

Moindres carrés dans un réseau de capteurs : cas vectoriel

TD3, ex. IV.2

Remarque:

- Dans la limite précédente:

$$\widehat{\boldsymbol{\theta}} = \lim_{k \to \infty} \mathbf{P}_i^{-1}(k)\, \widehat{\boldsymbol{\theta}}_i(k), \;\; i \in \{1, \ldots, n\}$$

  la matrice $\mathbf{P}_i(k)$ *n'est pas nécessairement inversible* pour toute valeur de k.

- Par conséquent, le capteur i peut calculer son estimé local de $\boldsymbol{\theta}$ seulement lorsque $\mathbf{P}_i(k)$ devient *non singulière*

Pour plus de détails, voir:

"*Distributed linear parameter estimation over wireless sensor networks*", A. Das, M. Mesbahi, IEEE Trans. Aerospace and Electronic Systems, vol. 45, n. 4, pp. 1293-1306, 2009