

---

# Résolution de systèmes linéaires Moindres carrés récursifs et Filtre de Kalman discret

---

Version 1.0

12 décembre 2008

Michel Libre

---

Ref. DCSD-2008\_069-NOT-001-1.0





# Table des matières

<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Inversion d'une matrice par bloc</b>	<b>7</b>
<b>3</b>	<b>Les lemmes d'inversions matricielles</b>	<b>9</b>
3.1	Premier lemme d'inversion matricielle . . . . .	9
3.2	Deuxième lemme d'inversion matricielle . . . . .	9
3.3	Une inversion itérative . . . . .	10
<b>4</b>	<b>Résolution des systèmes linéaires</b>	<b>11</b>
4.1	La solution des moindres carrés : système sur-déterminé régulier . . . . .	12
4.2	La pseudo inverse de Moore-Penrose : système sous-déterminé régulier . . . . .	12
4.3	Systèmes quelconques : . . . . .	14
4.3.1	La factorisation QRE . . . . .	14
4.3.2	La factorisation SVD : Décomposition en valeurs singulières . . . . .	15
4.4	L'inverse robuste singulière et le gain de Kalman . . . . .	16
4.4.1	Première forme - par l'inversion de la matrice d'information . . . . .	16
4.4.2	Deuxième forme - utilisant les matrices de covariance . . . . .	17
4.4.3	Comparaison des deux formes . . . . .	17
4.4.4	Forme scalaire de l'approche par la matrice d'information . . . . .	17
4.4.5	Forme scalaire l'approche par la matrice de covariance . . . . .	18
<b>5</b>	<b>Moindres carrés récursif et filtre de Kalman discret</b>	<b>19</b>
5.1	Estimation par moindres carrés classiques . . . . .	19
5.2	Première forme récursive . . . . .	20
5.3	Deuxième forme récursive . . . . .	21
5.4	Comparaison des deux formes récursives . . . . .	22
5.5	Liaison avec les formules du filtre de Kalman discret . . . . .	22
5.5.1	Covariance de l'erreur d'estimation dans le cas de moindres carrés récursifs . . . . .	22
5.5.2	Estimation de la variance des erreurs de mesure . . . . .	23
5.6	Compléments sur l'estimation . . . . .	24
5.6.1	La propagation de l'estimation . . . . .	24
5.6.2	La formule de Joseph . . . . .	25
5.6.3	Gain de Kalman en continu et en discret . . . . .	26

<b>6 Les bruits dans les filtres de Kalman</b>	<b>27</b>
6.1 Le bruit d'état . . . . .	27
6.1.1 Système continu . . . . .	27
6.1.2 Système discret . . . . .	28
6.1.3 Simulation numérique approchée du processus continu . . . . .	29
6.1.4 Conclusion sur les représentations de bruit d'état . . . . .	30
6.2 Le bruit de mesure . . . . .	30
6.2.1 Conclusion sur les représentations de bruit de mesure . . . . .	31
6.3 Exemple : le gyromètre . . . . .	32
6.4 Régimes permanents des filtres . . . . .	33
6.4.1 Filtre continu . . . . .	33
6.4.2 Filtre numérique . . . . .	33
6.5 Annexe . . . . .	34

# Chapitre 1

## Introduction

Cette note<sup>1</sup> présente des résultats pratiques pour aider à la mise en oeuvre de techniques d'estimations numériques. Les démonstrations qui sont proposées dans cette note n'ont aucune prétention en terme de rigueur mathématique. Elles sont données pour faciliter la compréhension et la maîtrise des techniques présentées.

Le premier chapitre n'est qu'un prétexte à l'obtention des lemmes d'inversion matricielles présentés au deuxième chapitre. Ce sont deux formules magiques qui sont fréquemment utilisées dans le reste de cette note. Il suffit de savoir que ces formules existent et où les trouver.

Le troisième chapitre présente la résolution des systèmes linéaires, avec prise en compte des divers cas particuliers (plus ou moins d'équations que d'inconnues, ...) et les méthodes associées.

Le quatrième chapitre présente la méthode des moindres carrés récursifs qui permet d'estimer des inconnues à partir de mesures linéaires en fonction de ces inconnues. Une des formes de la solution obtenue correspond aux équations du filtre de Kalman discret.

Le dernier chapitre précise les unités qui sont employées pour qualifier les bruits d'état et de mesure lors de la mise en oeuvre des filtres de Kalman continu et discret.

---

1. Un petit merci à Alexandre Amiez et à Claude Lambert pour la re-lecture de cette note.



## Chapitre 2

# Inversion d'une matrice par bloc

Soit à inverser une matrice carrée  $\mathbf{H}$  constituée de 4 blocs :

$$\mathbf{H} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{D} & \mathbf{C} \end{pmatrix}$$

avec  $\mathbf{A}$  régulière de dimension  $n \times n$ ,  $\mathbf{C}$  régulière de dimension  $m \times m$  et  $\mathbf{B}$  et  $\mathbf{D}$  quelconques de dimensions respectives  $n \times m$  et  $m \times n$ .  $\mathbf{H}$  est carrée régulière de dimension  $(n+m) \times (n+m)$ . Ecrivons l'inverse  $\mathbf{H}^{-1}$  sous la forme suivante :

$$\mathbf{H}^{-1} = \begin{pmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{W} & \mathbf{Z} \end{pmatrix}$$

avec  $\mathbf{X}$  de dimension  $n \times n$ ,  $\mathbf{Z}$  de dimension  $m \times m$  et  $\mathbf{Y}$  et  $\mathbf{W}$  de dimensions respectives  $n \times m$  et  $m \times n$ .

Effectuons le produit :

$$\mathbf{H}\mathbf{H}^{-1} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{D} & \mathbf{C} \end{pmatrix} \begin{pmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{W} & \mathbf{Z} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_{n \times n} & \mathbf{0}_{n \times m} \\ \mathbf{0}_{m \times n} & \mathbf{I}_{m \times m} \end{pmatrix}$$

Il vient :

$$\mathbf{A}\mathbf{X} + \mathbf{B}\mathbf{W} = \mathbf{I}_{n \times n} \quad (2.1)$$

$$\mathbf{A}\mathbf{Y} + \mathbf{B}\mathbf{Z} = \mathbf{0}_{n \times m} \quad (2.2)$$

$$\mathbf{D}\mathbf{X} + \mathbf{C}\mathbf{W} = \mathbf{0}_{m \times n} \quad (2.3)$$

$$\mathbf{D}\mathbf{Y} + \mathbf{C}\mathbf{Z} = \mathbf{I}_{m \times m} \quad (2.4)$$

Compte tenu de la régularité de  $\mathbf{A}$ , (2.2) implique :

$$\mathbf{Y} = -\mathbf{A}^{-1}\mathbf{B}\mathbf{Z}$$

qui reporté dans (2.4) implique :

$$(\mathbf{C} - \mathbf{D}\mathbf{A}^{-1}\mathbf{B})\mathbf{Z} = \mathbf{I}$$

soit :

$$\begin{aligned} \mathbf{Z} &= (\mathbf{C} - \mathbf{D}\mathbf{A}^{-1}\mathbf{B})^{-1} \\ \mathbf{Y} &= -\mathbf{A}^{-1}\mathbf{B}(\mathbf{C} - \mathbf{D}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{aligned} \quad (2.5)$$

Compte tenu de la régularité de  $\mathbf{C}$ , (2.3) implique :

$$\mathbf{W} = -\mathbf{C}^{-1}\mathbf{D}\mathbf{X}$$

qui reporté dans (2.1) implique :

$$(\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{D})\mathbf{X} = \mathbf{I}$$

soit :

$$\begin{aligned} \mathbf{X} &= (\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{D})^{-1} \\ \mathbf{W} &= -\mathbf{C}^{-1}\mathbf{D}(\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{D})^{-1} \end{aligned} \quad (2.6)$$

d'où :

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{D} & \mathbf{C} \end{pmatrix}^{-1} = \begin{pmatrix} (\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{D})^{-1} & -\mathbf{A}^{-1}\mathbf{B}(\mathbf{C} - \mathbf{D}\mathbf{A}^{-1}\mathbf{B})^{-1} \\ -\mathbf{C}^{-1}\mathbf{D}(\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{D})^{-1} & (\mathbf{C} - \mathbf{D}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{pmatrix}$$



## Chapitre 3

# Les lemmes d'inversions matricielles

Les deux lemmes d'inversion matricielles sont des formules qui lient des expressions matricielles que l'on rencontre souvent dans les problèmes d'optimisation.

Considérons la matrice  $\mathbf{H}$  précédente, mais symétrique, c'est-à-dire telle que :

$$\begin{aligned}\mathbf{A} &= \mathbf{A}^T \\ \mathbf{C} &= \mathbf{C}^T \\ \mathbf{B} &= \mathbf{D}^T\end{aligned}$$

Il en résulte  $\mathbf{H}^{-1}$  est symétrique, soit :

$$\begin{aligned}\mathbf{X} &= \mathbf{X}^T \\ \mathbf{Z} &= \mathbf{Z}^T \\ \mathbf{Y} &= \mathbf{W}^T\end{aligned}$$

Cette dernière relation s'écrit :

$$-\mathbf{A}^{-1}\mathbf{B}(\mathbf{C} - \mathbf{B}^T\mathbf{A}^{-1}\mathbf{B})^{-1} = -(\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{C}^{-1}$$

### 3.1 Premier lemme d'inversion matricielle

On obtient le *premier lemme d'inversion matricielle* en remplaçant  $\mathbf{A}^{-1}$  par  $-\mathbf{A}$  dans la relation précédente :

$$\boxed{\mathbf{A}\mathbf{B}(\mathbf{C} + \mathbf{B}^T\mathbf{A}\mathbf{B})^{-1} = (\mathbf{A}^{-1} + \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{C}^{-1}} \quad (3.1)$$

### 3.2 Deuxième lemme d'inversion matricielle

La relation (2.1) s'écrit :

$$\mathbf{X} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\mathbf{W}$$

soit en remplaçant  $\mathbf{W}$  par  $\mathbf{Y}^T$  tiré de (2.5) :

$$\mathbf{X} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{C} - \mathbf{B}^T\mathbf{A}^{-1}\mathbf{B})\mathbf{B}^T\mathbf{A}^{-1}$$

qui comparé à (2.6) donne :

$$\mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{C} - \mathbf{B}^T\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{A}^{-1} = (\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T)^{-1}$$

On obtient le *deuxième lemme d'inversion matricielle* en remplaçant  $\mathbf{A}^{-1}$  par  $-\mathbf{A}$  et en changeant tous les signes :

$$\boxed{(\mathbf{A}^{-1} + \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T)^{-1} = \mathbf{A} - \mathbf{A}\mathbf{B}(\mathbf{C} + \mathbf{B}^T\mathbf{A}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{A}} \quad (3.2)$$

### 3.3 Une inversion itérative

Le calcul de l'inverse de  $\mathbf{A}^{-1} + \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T$  se rencontre dans de nombreux problèmes d'optimisation. Ce calcul peut être réalisé par une procédure récursive simple, qui n'inverse que  $m$  scalaires dans le cas où  $\mathbf{C} = \Lambda$  est une matrice diagonale. Notons  $\lambda_i$  les éléments diagonaux de  $\mathbf{C}$  et  $\mathbf{b}_i$  les vecteurs colonnes de  $\mathbf{B}$ . On obtient alors :

$$\begin{aligned} \mathbf{A}^{-1} + \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T &= \mathbf{A}^{-1} + (\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_m) \begin{pmatrix} \frac{1}{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\lambda_m} \end{pmatrix} \begin{pmatrix} \mathbf{b}_1^T \\ \vdots \\ \mathbf{b}_m^T \end{pmatrix} \\ &= \mathbf{A}^{-1} + \begin{pmatrix} \frac{1}{\lambda_1}\mathbf{b}_1 & \cdots & \frac{1}{\lambda_m}\mathbf{b}_m \end{pmatrix} \begin{pmatrix} \mathbf{b}_1^T \\ \vdots \\ \mathbf{b}_m^T \end{pmatrix} \\ &= \mathbf{A}^{-1} + \frac{1}{\lambda_1}\mathbf{b}_1\mathbf{b}_1^T + \cdots + \frac{1}{\lambda_m}\mathbf{b}_m\mathbf{b}_m^T \end{aligned}$$

Notons  $\mathbf{A}_{k-1}$  et  $\mathbf{A}_k$  les matrices telles que :

$$\begin{aligned} \mathbf{A}_{k-1}^{-1} &= \mathbf{A}^{-1} + \frac{1}{\lambda_1}\mathbf{b}_1\mathbf{b}_1^T + \cdots + \frac{1}{\lambda_{k-1}}\mathbf{b}_{k-1}\mathbf{b}_{k-1}^T \\ \mathbf{A}_k^{-1} &= \mathbf{A}^{-1} + \frac{1}{\lambda_1}\mathbf{b}_1\mathbf{b}_1^T + \cdots + \frac{1}{\lambda_k}\mathbf{b}_k\mathbf{b}_k^T \end{aligned}$$

soit :

$$\mathbf{A}_k^{-1} = \mathbf{A}_{k-1}^{-1} + \frac{1}{\lambda_k}\mathbf{b}_k\mathbf{b}_k^T$$

et calculons  $\mathbf{A}_k$  connaissant  $\mathbf{A}_{k-1}$ . En appliquant le deuxième lemme d'inversion matricielle, il vient :

$$\mathbf{A}_k = \mathbf{A}_{k-1} - \mathbf{A}_{k-1}\mathbf{b}_k(\lambda_k + \mathbf{b}_k^T\mathbf{A}_{k-1}\mathbf{b}_k)^{-1}\mathbf{b}_k^T\mathbf{A}_{k-1}$$

soit :

$$\mathbf{A}_k = \mathbf{A}_{k-1} - \frac{1}{\lambda_k + \mathbf{b}_k^T\mathbf{A}_{k-1}\mathbf{b}_k}\mathbf{A}_{k-1}\mathbf{b}_k\mathbf{b}_k^T\mathbf{A}_{k-1} \quad (3.3)$$

Le calcul de  $\mathbf{A}_m = (\mathbf{A}^{-1} + \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T)^{-1} = \mathbf{A} - \mathbf{A}\mathbf{B}(\mathbf{C} + \mathbf{B}^T\mathbf{A}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{A}$  peut être fait en appliquant la formule de récurrence (3.3)  $m$  fois avec l'initialisation  $\mathbf{A}_0 = \mathbf{A}$ .

Si  $\mathbf{A}$  est symétrique, alors toutes les matrices  $\mathbf{A}_k$  le sont. Le calcul de  $\mathbf{A}_m$  se simplifie en :

$$\mathbf{A}_0 = \mathbf{A} \text{ puis } \begin{cases} \mathbf{v}_k = \mathbf{A}_{k-1}\mathbf{b}_k \\ \mathbf{A}_k = \mathbf{A}_{k-1} - \frac{1}{\lambda_k + \mathbf{b}_k^T\mathbf{v}_k}\mathbf{v}_k\mathbf{v}_k^T \end{cases} \text{ pour } k = 1 \text{ à } m.$$

## Chapitre 4

# Résolution des systèmes linéaires

Considérons la résolution du système linéaire :

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

où  $\mathbf{x}$  et  $\mathbf{y}$  sont deux vecteurs respectivement de dimension  $n$  et  $m$  et  $\mathbf{A}$  une matrice  $m \times n$  de rang  $r$ .

$$\begin{aligned} n &= \dim(\mathbf{x}) \\ m &= \dim(\mathbf{y}) \\ r &= \text{rang}(\mathbf{A}) \end{aligned}$$

Nous dirons que ce système est :

- déterminé si  $m = n$ ,
- déterminé régulier si  $m = n = r$
- sur-déterminé si  $m > n$
- sur-déterminé régulier si  $m > n = r$
- sous-déterminé ou redondant si  $m < n$
- redondant régulier si  $r = m < n$

Si le système est déterminé régulier on écrit la solution sous la forme classique :

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$$

D'un point de vue pratique, on calcule directement  $\mathbf{x}$  à l'aide d'un des algorithmes suivants :

- résolution directe par substitution si  $\mathbf{A}$  est triangulaire,
- factorisation LU (Doolittle ou Crout) si  $\mathbf{A}$  est quelconque,
- factorisation de Cholesky (de préférence sans racine carrée) si  $\mathbf{A}$  est symétrique.

Si le système est sur-déterminé régulier, on utilise en général la solution approchée des moindres carrés décrite ci-après.

Si le système est sous-déterminé régulier, on utilise en général la solution exacte de Moore-Penrose décrite ensuite.

Les autres cas sont plus complexes. Des solutions approchées sont fournies en utilisant la factorisation QRE, la décomposition en valeur singulière ou encore l'inverse robuste singulière que nous présentons plus loin dans ce document.

## 4.1 La solution des moindres carrés : système sur-déterminé régulier

Considérons le cas des systèmes réguliers ayant plus de conditions indépendantes que d'inconnues. Ils sont tels que :

$$m = \dim(\mathbf{y}) > \dim(\mathbf{x}) = n = \text{rang}(\mathbf{A})$$

On suppose donc que la matrice  $\mathbf{A}$  est de rang complet. Dans ce cas, il n'y a, a priori, pas de solution. On cherche malgré tout le vecteur  $\mathbf{x}$  qui produit une erreur  $\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{A}\mathbf{x}$  minimale. D'où le critère :

$$r = \frac{1}{2} (\mathbf{y} - \mathbf{A}\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{A}\mathbf{x})$$

où  $\mathbf{R}$  est une matrice symétrique<sup>1</sup> régulière de pondérations.

$$\frac{\partial r}{\partial \mathbf{x}} = -\mathbf{A}^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{A}\mathbf{x}) = 0 \rightarrow \mathbf{x} = (\mathbf{A}^T \mathbf{R}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{R}^{-1} \mathbf{y}$$

Cette solution est surtout employée avec  $\mathbf{R}^{-1} = \mathbf{1}$ , d'où la pseudo-inverse classique des moindres carrés :

$$\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y}$$

On remarquera que cette solution inverse une matrice  $n \times n$  (la plus petite des dimensions de  $\mathbf{A}$ ). Elle n'existe que si  $\mathbf{A}$  est de rang complet.

On peut calculer  $\mathbf{x}$  à l'aide de la factorisation de Cholesky. On évaluera d'abord le vecteur  $\mathbf{z} = \mathbf{A}^T \mathbf{R}^{-1} \mathbf{y}$ , puis on résoudra  $(\mathbf{A}^T \mathbf{R}^{-1} \mathbf{A}) \mathbf{x} = \mathbf{z}$  en utilisant la factorisation de Cholesky.

### Interprétation géométrique :

Dans la solution classique des moindres carrés,  $\mathbf{x}$  est calculé à partir de la relation  $\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{y}$ . Dans le premier membre de cette relation le terme  $\mathbf{A} \mathbf{x} = \mathbf{v}$  représente un vecteur  $\mathbf{v}$  combinaison linéaire des vecteurs colonnes de la matrice  $\mathbf{A}$  dont les  $x_i$  sont les composantes contravariantes. Ensuite  $\mathbf{A}^T \mathbf{v} = \mathbf{A}^T \mathbf{y}$  exprime le fait que les projections de ce vecteur  $\mathbf{v}$  sur ces vecteurs colonnes (les composantes covariantes) sont identiques aux projections de  $\mathbf{y}$  sur ces colonnes. On voit donc que seule la projection de  $\mathbf{y}$  sur le sous-espace image de  $\mathbf{A}$  (l'espace engendré par ses vecteurs colonnes) est restituée par la solution  $\mathbf{A} \mathbf{x}$ .

L'introduction de la matrice de pondération  $\mathbf{M} = \mathbf{R}^{-1}$ , modifie la relation précédente en  $\mathbf{A}^T \mathbf{M} \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{M} \mathbf{y}$ . Ce qui fait que les projections sur les colonnes de  $\mathbf{A}$  qui sont rendues identiques sont celles de  $\mathbf{M} \mathbf{y}$  avec celles de  $\mathbf{M} \mathbf{v}$ .

## 4.2 La pseudo inverse de Moore-Penrose : système sous-déterminé régulier

Considérons le cas des systèmes réguliers ayant plus d'inconnues que d'équations. Ils sont tels que :

$$\text{rang}(\mathbf{A}) = m = \dim(\mathbf{y}) \leq \dim(\mathbf{x}) = n$$

On suppose donc que la matrice  $\mathbf{A}$  est de rang complet. Si  $m < n$  (strictement), il y a une infinité de solutions  $\mathbf{x}$ . Dans cette infinité on choisit celle qui minimise la norme  $\|\mathbf{x} - \mathbf{x}_0\|$  avec généralement  $\mathbf{x}_0 = 0$ . D'où le critère à minimiser :

$$r = \frac{1}{2} (\mathbf{x} - \mathbf{x}_0)^T \mathbf{P}^{-1} (\mathbf{x} - \mathbf{x}_0) + (\mathbf{y} - \mathbf{A}\mathbf{x})^T \boldsymbol{\lambda}$$

1. Car la forme quadratique  $r$  est un scalaire tel que  $2r = \boldsymbol{\varepsilon}^T \mathbf{M} \boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^T \mathbf{M}^T \boldsymbol{\varepsilon} = \frac{1}{2} \boldsymbol{\varepsilon}^T (\mathbf{M} + \mathbf{M}^T) \boldsymbol{\varepsilon}$ . On peut donc toujours considérer que  $\mathbf{R}^{-1} = \frac{1}{2} (\mathbf{M} + \mathbf{M}^T)$  est symétrique

$\lambda$  est le vecteur des multiplicateurs de Lagrange associés aux contraintes  $\mathbf{y} - \mathbf{A}\mathbf{x} = \mathbf{0}$  et  $\mathbf{P}^{-1}$  est une matrice de pondérations symétrique régulière choisie pour favoriser la proximité de certaines composantes par rapport à celle des autres. Les conditions de stationnarité du premier ordre donnent :

$$\begin{aligned}\frac{\partial r}{\partial \mathbf{x}} &= \mathbf{P}^{-1}(\mathbf{x} - \mathbf{x}_0) - \mathbf{A}^T \lambda = 0 \rightarrow \mathbf{x} = \mathbf{x}_0 + \mathbf{P}\mathbf{A}^T \lambda \\ \mathbf{y} &= \mathbf{A}\mathbf{x}_0 + \mathbf{A}\mathbf{P}\mathbf{A}^T \lambda \rightarrow \lambda = (\mathbf{A}\mathbf{P}\mathbf{A}^T)^{-1}(\mathbf{y} - \mathbf{A}\mathbf{x}_0)\end{aligned}$$

d'où :

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{P}\mathbf{A}^T (\mathbf{A}\mathbf{P}\mathbf{A}^T)^{-1}(\mathbf{y} - \mathbf{A}\mathbf{x}_0)$$

On remarquera que cette solution inverse une matrice  $\dim(\mathbf{y}) \times \dim(\mathbf{y})$  (la plus petite des dimensions de  $\mathbf{A}$ ).

Cette solution est surtout employée avec  $\mathbf{P} = \mathbf{1}$  et  $\mathbf{x}_0 = \mathbf{0}$ , ce qui correspond à rechercher *la solution de norme minimale*. Elle est fournie par la pseudo-inverse dite de Moore-Penrose  $\mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1}$  :

$$\mathbf{x} = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{y}$$

Elle n'existe que si  $\mathbf{A}$  est de rang complet (cf. hypothèse).

#### Interprétation géométrique :

Un raisonnement géométrique permet d'obtenir plus rapidement ce résultat. Décomposons  $\mathbf{x}$  en deux parties  $\mathbf{x}_e$  et  $\mathbf{x}_n$  :

$$\mathbf{x} = \mathbf{x}_e + \mathbf{x}_n \quad \text{avec} \quad \begin{cases} \mathbf{x}_e = \mathbf{A}^T \lambda \\ \mathbf{A}\mathbf{x}_n = 0 \end{cases}$$

$\mathbf{x}_e$  appartient aux sous-espace engendré par les lignes de  $\mathbf{A}$ , et  $\mathbf{x}_n$  (dans le noyau de  $\mathbf{A}$ ) est orthogonal à ce sous-espace. Il en résulte que :

$$\mathbf{y} = \mathbf{A}\mathbf{x} = \mathbf{A}(\mathbf{x}_e + \mathbf{x}_n) = \mathbf{A}\mathbf{A}^T \lambda$$

Si  $\mathbf{A}$  est de rang complet, alors  $\lambda = (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{y}$  d'où :

$$\mathbf{x}_e = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{y}$$

La solution générale consiste à ajouter à la partie  $\mathbf{x}_e$  produisant effectivement  $\mathbf{y}$ , un vecteur  $\mathbf{x}_n$  dans le noyau de  $\mathbf{A}$  dite solution homogène ou solution nulle car elle ne produit rien au niveau de  $\mathbf{y}$ . Si on note :

$$\mathbf{A}^\dagger = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1}$$

la pseudo-inverse de Moore-Penrose, la matrice  $\mathbf{A}^\dagger \mathbf{A}$  est un projecteur orthogonal dans le sous-espace des lignes de  $\mathbf{A}$  et la matrice :

$$\mathbf{P}_N = \mathbf{1} - \mathbf{A}^\dagger \mathbf{A}$$

est un projecteur dans le noyau de  $\mathbf{A}$ , d'où quelque soit le vecteur  $\mathbf{x}'$  de l'espace des  $\mathbf{x}$  :

$$\mathbf{x}_n = (\mathbf{1} - \mathbf{A}^\dagger \mathbf{A}) \mathbf{x}'$$

Remarque : Pour résoudre ces systèmes, qui font toujours intervenir des matrices symétriques, on utilise souvent la factorisation de Cholesky.

### 4.3 Systèmes quelconques :

#### 4.3.1 La factorisation QRE

Dans le pire des cas, le vecteur  $\mathbf{y}$  n'appartient pas à l'espace engendré par les colonnes de  $\mathbf{A}$ . Le système  $\mathbf{y} = \mathbf{A}\mathbf{x}$  n'est pas compatible. Seule la projection  $\mathbf{y}_A$  de  $\mathbf{y}$  dans l'espace image de  $\mathbf{A}$  est réalisable et le système  $\mathbf{y}_A = \mathbf{A}\mathbf{x}$  a une infinité de solutions. La factorisation QRE, variante de la factorisation QR, nous semble la méthode la mieux adaptée à la résolution de ce problème. Elle peut même fournir la solution de norme minimale (au prix de deux factorisations successives) pour laquelle on fait généralement appel à la décomposition en valeurs singulières. La factorisation QRE permet également de déterminer le rang de la matrice  $\mathbf{A}$ .

L'algorithme de factorisation QRE met la matrice  $\mathbf{A}$  de dimension  $m \times n$  sous la forme :

$$\mathbf{A} = \mathbf{Q}\mathbf{R}\mathbf{E}^T \Leftrightarrow \mathbf{R} = \mathbf{Q}^T\mathbf{A}\mathbf{E}$$

où  $\mathbf{Q}$  est une matrice unitaire de dimension  $m \times m$  ( $\mathbf{Q}^{-1} = \mathbf{Q}^T$ ) et  $\mathbf{R}$  une matrice triangulaire supérieure de dimension  $m \times n$ , avec forcément ses  $m - n$  dernières lignes nulles.

$$\mathbf{R} = \begin{bmatrix} \mathbf{U} \\ \mathbf{0} \end{bmatrix}$$

et où  $\mathbf{E}$  est une matrice unitaire produit de matrices de permutations.

La résolution de  $\mathbf{y} = \mathbf{A}\mathbf{x} = \mathbf{Q}\mathbf{R}\mathbf{E}^T\mathbf{x}$  consiste à :

1. calculer  $\mathbf{z} = \mathbf{Q}^T\mathbf{y} = \mathbf{S}_r \dots \mathbf{S}_2\mathbf{S}_1\mathbf{y}$ ,
2. calculer  $\mathbf{w}_r$  solution de  $\mathbf{U}_r\mathbf{w}_r = \mathbf{z}_r$  où  $\mathbf{z}_r$  ne comporte que les  $r$  premières composantes de  $\mathbf{z}$ ,
3. former  $\mathbf{w}$  à partir de  $\mathbf{w}_r$  en ajoutant  $n - r$  dernières composantes nulles,
4. calculer  $\mathbf{x} = \mathbf{E}_1\mathbf{E}_2 \dots \mathbf{E}_r\mathbf{w}$ .

Dans le cas d'une simple résolution,  $\mathbf{z}$  est calculé au fur et à mesure en lieu et place de  $\mathbf{y}$  par les produits par  $\mathbf{S}_k$ .  $\mathbf{R}$  est calculée en lieu et place de  $\mathbf{A}$ .  $\mathbf{w}$  est calculé en lieu et place de  $\mathbf{z}$  (c'est-à-dire de  $\mathbf{y}$ ). Seuls les indices des permutations effectuées sont mémorisés et  $\mathbf{x}$  est directement obtenu en appliquant la permutation aux composantes de  $\mathbf{w}$ .

**Remarque :** Si  $r < \min(n, m)$ , la solution trouvée n'est pas de norme minimale, mais elle est exacte au sens que  $\mathbf{A}\mathbf{x}$  est égal à la partie réalisable de  $\mathbf{y}$  (la projection de  $\mathbf{y}$  dans l'espace image de  $\mathbf{A}^T$ ).

**Noyaux** de  $\mathbf{A}$  et  $\mathbf{A}^T$ .

Notons  $\mathbf{Q}_r$  les  $r$  premières colonnes de  $\mathbf{Q}$  et  $\mathbf{N}_g$  les  $m - r$  dernières dans la décomposition QRE de  $\mathbf{A}$  :

$$\mathbf{Q} = ( \mathbf{Q}_r \quad \mathbf{N}_g )$$

Il vient :

$$\begin{pmatrix} \mathbf{Q}_r^T \\ \mathbf{N}_g^T \end{pmatrix} \mathbf{A} = \begin{pmatrix} \mathbf{U}_r & \mathbf{B} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{E}^T$$

qui montre que :

$$\mathbf{N}_g^T \mathbf{A} = \mathbf{0} \quad ; \quad \mathbf{A}^T \mathbf{N}_g = \mathbf{0}$$

Les  $m - r$  dernières colonnes de  $\mathbf{Q}$  constituent une base orthonormée du noyau de  $\mathbf{A}^T$ .

Pour obtenir la solution de norme minimale de  $\mathbf{y} = \mathbf{A}\mathbf{x}$ , il faut retrancher à  $\mathbf{x}$  sa projection sur le noyau de  $\mathbf{A}$ . Si  $\mathbf{N}$  représente le noyau orthonormé de  $\mathbf{A}$ , la solution de norme minimale s'écrit :

$$\hat{\mathbf{x}} = \mathbf{x} - \mathbf{N}\mathbf{N}^T\mathbf{x} \quad (4.1)$$

Dans la résolution précédente, avec  $r < \min(n, m)$ , le noyau nécessaire pour calculer la solution optimale est celui de  $\mathbf{A}$ , alors qu'on dispose dans  $\mathbf{Q}$  de celui de  $\mathbf{A}^T$ . Il faut donc effectuer en plus une décomposition QRE de  $\mathbf{A}^T$ . Cette méthode peut s'avérer plus rapide que la méthode généralement utilisée qui passe par la décomposition en valeurs singulières de  $\mathbf{A}$ .

### 4.3.2 La factorisation SVD : Décomposition en valeurs singulières

La décomposition en valeurs singulières de  $\mathbf{A}$  s'écrit :

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T$$

où  $\mathbf{U}$  est une matrice unitaire  $m \times m$ ,  $\mathbf{V}$  une matrice unitaire  $n \times n$  et  $\mathbf{D}$  une matrice  $m \times n$  de la forme :

$$\mathbf{D} = \begin{pmatrix} \Lambda & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(m-r) \times r} & \mathbf{0}_{(m-r) \times (n-r)} \end{pmatrix}$$

où  $\Lambda$  est une matrice diagonale de dimension  $r \times r$  dont les éléments diagonaux sont les valeurs singulières de la matrice  $\mathbf{A}$ , généralement classées de la plus grande en haut à gauche, à la plus petite en bas à droite. On peut vérifier que :

$$\begin{aligned} (\mathbf{A}^T \mathbf{A}) \mathbf{V} &= \mathbf{V} \begin{pmatrix} \Lambda^2 & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{0}_{(n-r) \times (n-r)} \end{pmatrix} \\ (\mathbf{A} \mathbf{A}^T) \mathbf{U} &= \mathbf{U} \begin{pmatrix} \Lambda^2 & \mathbf{0}_{r \times (m-r)} \\ \mathbf{0}_{(m-r) \times r} & \mathbf{0}_{(m-r) \times (m-r)} \end{pmatrix} \end{aligned}$$

ce qui montre que les valeurs singulières sont les racines carrées des valeurs propres de  $\mathbf{A}^T \mathbf{A}$  et de  $\mathbf{A} \mathbf{A}^T$ , que  $\mathbf{U}$  est la matrice des vecteurs propres de  $\mathbf{A} \mathbf{A}^T$  et que  $\mathbf{V}$  est la matrice des vecteurs propres de  $\mathbf{A}^T \mathbf{A}$ .

La pseudo-inverse qui fournit la solution de norme minimale au système  $\mathbf{y} = \mathbf{A}\mathbf{x}$  s'écrit :

$$\begin{aligned} \mathbf{A}^\dagger &= \mathbf{V} \begin{pmatrix} \Lambda^{-1} & \mathbf{0}_{r \times (m-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{0}_{(n-r) \times (m-r)} \end{pmatrix} \mathbf{U}^T \\ \hat{\mathbf{x}} &= \mathbf{A}^\dagger \mathbf{y} \end{aligned}$$

En partitionnant la matrice  $\mathbf{V}$  en  $\mathbf{V} = (\mathbf{V}_r \quad \mathbf{N})$  avec  $\mathbf{V}_r$  de dimension  $n \times r$ , la matrice  $\mathbf{N}$  n'est autre que le noyau orthonormé de  $\mathbf{A}$  :

$$\mathbf{V} = (\mathbf{V}_r \quad \mathbf{N}) \text{ tel que } \mathbf{A}\mathbf{N} = \mathbf{0}$$

La solution générale s'écrit alors :

$$\mathbf{x} = \hat{\mathbf{x}} + \mathbf{x}_N \text{ avec } \begin{cases} \hat{\mathbf{x}} = \mathbf{A}^\dagger \mathbf{y} \text{ de norme minimale} \\ \mathbf{x}_N = \mathbf{N}\mathbf{N}^T \mathbf{x}', \forall \mathbf{x}' \in \mathbb{R}^n \end{cases}$$

Remarque :  $\mathbf{x}_N$  s'écrit également  $\mathbf{x}_N = \mathbf{N}\mathbf{z}$  où  $\mathbf{z}$  est un vecteur quelconque de dimension  $\in \mathbb{R}^{n-r}$ .

#### Noyau de $\mathbf{A}$

La factorisation SVD d'une matrice  $\mathbf{A}$  contient les noyaux à droite et à gauche de la matrice  $\mathbf{A}$ . Notons  $\mathbf{U}_r$  les  $r$  premières colonnes de  $\mathbf{U}$ ,  $\mathbf{N}_g$  les  $m-r$  dernières,  $\mathbf{V}_r$  les  $r$  premières colonnes de  $\mathbf{V}$  et  $\mathbf{N}_d$  les  $n-r$  dernières. Il vient :

$$\mathbf{A} = (\mathbf{U}_r \quad \mathbf{N}_g) \begin{pmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{V}_r^T \\ \mathbf{N}_d^T \end{pmatrix}$$

Ce qui montre que :

$$\mathbf{A}\mathbf{N}_d = \mathbf{0} \text{ et } \mathbf{N}_g^T \mathbf{A} = \mathbf{0}$$

Les  $n - r$  dernières colonnes de  $\mathbf{V}$  constituent une base orthonormée du noyau de  $\mathbf{A}$  et les  $m - r$  dernières colonnes de  $\mathbf{U}$  constituent une base orthonormée du noyau de  $\mathbf{A}^T$ .

**Attention :** Contrairement aux autres méthodes, la décomposition en valeurs singulières est un algorithme itératif censé converger vers les matrices recherchées. Le nombre d'opérations dépend de la précision recherchée.

## 4.4 L'inverse robuste singulière et le gain de Kalman

La presque-inverse, plus connue sous l'appellation inverse robuste singulière, fournit la solution  $\hat{\mathbf{x}}$  qui minimise un critère pondéré entre :

- la réalisation de l'égalité  $\mathbf{x} = \mathbf{x}_d$  avec une pondération quadratique  $\mathbf{P}_d^{-1}$ ,
- la réalisation de l'égalité  $\mathbf{y} = \mathbf{A}\mathbf{x}$  avec une pondération quadratique  $\mathbf{r}^{-1}$ .

Le critère pondéré s'écrit :

$$r = \frac{1}{2} (\mathbf{y} - \mathbf{A}\mathbf{x})^T \mathbf{r}^{-1} (\mathbf{y} - \mathbf{A}\mathbf{x}) + \frac{1}{2} (\mathbf{x} - \mathbf{x}_d)^T \mathbf{P}_d^{-1} (\mathbf{x} - \mathbf{x}_d)$$

### 4.4.1 Première forme - par l'inversion de la matrice d'information

La condition de stationnarité du premier ordre s'écrit :

$$\frac{\partial r}{\partial \mathbf{x}} = -\mathbf{A}^T \mathbf{r}^{-1} (\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}) + \mathbf{P}_d^{-1} (\hat{\mathbf{x}} - \mathbf{x}_d) = \mathbf{0}$$

soit :

$$(\mathbf{A}^T \mathbf{r}^{-1} \mathbf{A} + \mathbf{P}_d^{-1}) \hat{\mathbf{x}} - (\mathbf{A}^T \mathbf{r}^{-1} \mathbf{y} + \mathbf{P}_d^{-1} \mathbf{x}_d) = \mathbf{0} \quad (4.2)$$

D'où :

$$\hat{\mathbf{x}} = (\mathbf{A}^T \mathbf{r}^{-1} \mathbf{A} + \mathbf{P}_d^{-1})^{-1} (\mathbf{A}^T \mathbf{r}^{-1} \mathbf{y} + \mathbf{P}_d^{-1} \mathbf{x}_d) \quad (4.3)$$

Généralement on choisit  $\mathbf{r}^{-1} = \mathbf{1}$  pour que  $\mathbf{y}$  soit très proche de  $\mathbf{A}\mathbf{x}$ , et  $\mathbf{P}_d^{-1} = \varepsilon \mathbf{1}$  avec  $\varepsilon$  petit pour attirer la solution  $\hat{\mathbf{x}}$  vers la position favorite  $\mathbf{x}_d$ . Même si  $\mathbf{A}$  n'est pas de rang complet, il suffit de trouver une valeur minimale de  $\varepsilon$  qui rende  $\mathbf{A}^T \mathbf{A} + \varepsilon \mathbf{1}$  régulière (de rang  $n$ ).

Cette solution correspond à la formule de correction des moindres carrés récursifs (chapitre suivant) ou du *filtre de Kalman*.  $\hat{\mathbf{x}}$  représente la meilleure estimée d'un état inconnu  $\mathbf{x}$  dont on a une estimation a priori  $\mathbf{x}_d$  non biaisée, avec une certaine covariance d'erreur d'estimation  $\mathbf{P}_d$ .  $\mathbf{y}$  représente une mesure ultérieure qui est théoriquement reliée à  $\mathbf{x}$  par la relation  $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}$  où  $\mathbf{w}$  est un bruit de mesure à moyenne nulle de covariance  $\mathbf{r}$ .

Dans le cas particulier où  $\mathbf{A}$  est une matrice identité, c'est-à-dire si  $\mathbf{y}$  est une mesure directe de  $\mathbf{x}$ , cette solution s'écrit :

$$(\mathbf{r}^{-1} + \mathbf{P}_d^{-1}) \hat{\mathbf{x}} = \mathbf{P}_d^{-1} \mathbf{x}_d + \mathbf{r}^{-1} \mathbf{y}$$

Cette relation permet d'interpréter  $\hat{\mathbf{x}}$  comme barycentre de son estimée  $\mathbf{x}_d$  affecté de la pondération  $\mathbf{P}_d^{-1}$  et de sa mesure directe  $\mathbf{y}$  affectée de la pondération  $\mathbf{r}^{-1}$ . Ces matrices de pondérations s'interprètent comme les pertinences de ces deux informations. La pertinence se comporte ainsi de manière additive. En posant  $\mathbf{P}_{\hat{\mathbf{x}}}^{-1} = \mathbf{r}^{-1} + \mathbf{P}_d^{-1}$  la pertinence  $\mathbf{P}_{\hat{\mathbf{x}}}^{-1}$  de la nouvelle information  $\hat{\mathbf{x}}$  est la somme des deux pertinences utilisées, et dans le cas général où  $\mathbf{A}$  n'est pas unitaire, la nouvelle pertinence est le facteur de  $\hat{\mathbf{x}}$  dans l'expression (4.2), soit :

$$\mathbf{P}_{\hat{\mathbf{x}}}^{-1} = \mathbf{P}_d^{-1} + \mathbf{A}^T \mathbf{r}^{-1} \mathbf{A} \quad (4.4)$$



#### 4.4.2 Deuxième forme - utilisant les matrices de covariance

En appliquant le premier lemme d'inversion matricielle à (4.3) on obtient :

$$\hat{\mathbf{x}} = \mathbf{x}_d + \mathbf{P}_d \mathbf{A}^T (\mathbf{r} + \mathbf{A} \mathbf{P}_d \mathbf{A}^T)^{-1} (\mathbf{y} - \mathbf{A} \mathbf{x}_d) \quad (4.5)$$

qui met en évidence le gain de Kalman :

$$\mathbf{K} = \mathbf{P}_d \mathbf{A}^T (\mathbf{r} + \mathbf{A} \mathbf{P}_d \mathbf{A}^T)^{-1} \quad (4.6)$$

En utilisant le deuxième lemme d'inversion matricielle sur (4.4), on obtient la relation :

$$\begin{aligned} \mathbf{P}_{\hat{\mathbf{x}}} &= \mathbf{P}_d - \mathbf{P}_d \mathbf{A}^T (\mathbf{r} + \mathbf{A} \mathbf{P}_d \mathbf{A}^T)^{-1} \mathbf{A} \mathbf{P}_d \\ \mathbf{P}_{\hat{\mathbf{x}}} &= \mathbf{P}_d - \mathbf{K} \mathbf{A} \mathbf{P}_d \end{aligned} \quad (4.7)$$

qui est la formule de mise à jour de la matrice de covariance.

#### 4.4.3 Comparaison des deux formes

Sous sa première forme, la solution inverse une matrice  $n \times n$ , alors que sous la deuxième elle inverse une matrice  $m \times m$ . On choisira la forme qui inverse la plus petite matrice.

Si  $m < n$ , et si  $\text{rang}(\mathbf{A}) = m$ , la solution trouvée est souvent la même que celle fournie par la pseudo-inverse de Moore-Penrose.

La deuxième forme inverse une matrice de même dimension que la solution de Moore-Penrose. Elle peut être plus intéressante s'il n'est pas certain que le  $\text{rang}(\mathbf{A})$  soit toujours égal à  $m$ . En se reportant à la solution de Moore-Penrose, on constate qu'elles ne diffèrent que par la matrice  $\mathbf{r}$  qui est ajoutée  $\mathbf{A} \mathbf{P}_d \mathbf{A}^T$ . C'est cet ajout qui permet l'inversion lorsque le rang de  $\mathbf{A}$  devient inférieur à  $m$ .

#### 4.4.4 Forme scalaire de l'approche par la matrice d'information

Si la matrice  $\mathbf{r}^{-1}$  est diagonale, la solution  $\mathbf{x} = \mathbf{P}_{\hat{\mathbf{x}}} (\mathbf{A}^T \mathbf{r}^{-1} \mathbf{y} + \mathbf{P}_d^{-1} \mathbf{x}_d)$  peut être calculée sans inversion de matrice, en utilisant la méthode de l'inversion itérative décrite en 3.3 pour calculer  $\mathbf{P}_{\hat{\mathbf{x}}} = (\mathbf{A}^T \mathbf{r}^{-1} \mathbf{A} + \mathbf{P}_d^{-1})^{-1}$ . En notant  $r_i$  les éléments diagonaux de  $\mathbf{r}$  et  $\mathbf{a}_i$  les vecteurs colonnes  $\mathbf{A}^T$ , il vient :

$$\mathbf{x} = \left( \mathbf{P}_d^{-1} + \frac{1}{r_1} \mathbf{a}_1 \mathbf{a}_1^T + \dots + \frac{1}{r_m} \mathbf{a}_m \mathbf{a}_m^T \right)^{-1} \left( \mathbf{P}_d^{-1} \mathbf{x}_d + \frac{y_1}{r_1} \mathbf{a}_1 + \dots + \frac{y_m}{r_m} \mathbf{a}_m \right)$$

En posant  $\mathbf{P}_0 = \mathbf{P}_d$  et  $\mathbf{P}_m = \mathbf{P}_{\hat{\mathbf{x}}}$ ,  $\mathbf{P}_{\hat{\mathbf{x}}}$  est obtenu par les  $m$  itérations :

$$\begin{aligned} \mathbf{v}_k &= \mathbf{P}_{k-1} \mathbf{a}_k \\ \mathbf{P}_k &= \mathbf{P}_{k-1} - \frac{1}{r_k + \mathbf{a}_k^T \mathbf{v}_k} \mathbf{v}_k \mathbf{v}_k^T \end{aligned}$$

A chaque itération, on peut éventuellement calculer une estimation  $\hat{\mathbf{x}}_k$  de  $\mathbf{x}$  par :

$$\begin{aligned} \mathbf{s}_k &= \mathbf{s}_{k-1} + \frac{y_1}{r_1} \mathbf{a}_1 \\ \hat{\mathbf{x}}_k &= \mathbf{P}_k \mathbf{s}_k \end{aligned}$$

avec l'initialisation  $\mathbf{s}_0 = \mathbf{P}_0^{-1} \hat{\mathbf{x}}_0$ . A la fin, on a  $\hat{\mathbf{x}} = \hat{\mathbf{x}}_m$ . On peut tout aussi bien calculer directement  $\hat{\mathbf{x}} = \mathbf{P}_m \mathbf{s}_m$  sans calculer les  $\hat{\mathbf{x}}_k$  intermédiaires.

Cette approche à l'inconvénient d'utiliser les matrices  $\mathbf{P}_d$  et  $\mathbf{P}_d^{-1}$ . Ainsi, si on reçoit un nouveau vecteur de mesures  $\mathbf{y}$  et qu'on désire itérer le processus à partir du  $\hat{\mathbf{x}}$  que l'on vient de calculer (qui va jouer le rôle de  $\mathbf{x}_d$ ),  $\mathbf{P}_{\hat{\mathbf{x}}}$  va jouer le rôle de  $\mathbf{P}_d$ , il faudra l'inverser pour obtenir  $\mathbf{P}_d^{-1}$ .

#### 4.4.5 Forme scalaire l'approche par la matrice de covariance

Toujours dans le cas où  $\mathbf{r}$  est diagonale d'éléments diagonaux  $r_i$ , on traite les mesures  $y_i$  une par une. On calcule ainsi :

$$\begin{cases} \mathbf{k}_i = \frac{1}{r_i + \mathbf{a}_i^T \mathbf{P}_{i-1} \mathbf{a}_i} \mathbf{P}_{i-1} \mathbf{a}_i \\ \hat{\mathbf{x}}_i = \hat{\mathbf{x}}_{i-1} + \mathbf{k}_i (y_i - \mathbf{a}_i^T \mathbf{x}_{i-1}) \\ \mathbf{P}_i = \mathbf{P}_{i-1} - \mathbf{k}_i \mathbf{a}_i^T \mathbf{P}_{i-1} \end{cases} \quad \text{pour } i = 1 \text{ à } m$$

que l'on initialise avec  $\mathbf{P}_0 = \mathbf{P}_d$ . A l'issue des  $m$  itération, on a  $\hat{\mathbf{x}} = \hat{\mathbf{x}}_m$  et  $\mathbf{P}_{\hat{\mathbf{x}}} = \mathbf{P}_m$ .

Cette approche à l'avantage de fournir une matrice  $\mathbf{P}_{\hat{\mathbf{x}}}$  de même nature que la matrice initiale utilisée  $\mathbf{P}_d$ . On peut ainsi traiter l'arrivée de nouvelles mesures, sans avoir d'inversion à réaliser.

## Chapitre 5

# Moindres carrés récursif et filtre de Kalman discret

Soit à estimer un vecteur inconnu  $\mathbf{x}$  de dimension  $n$  à partir de  $k$  lots de vecteurs de mesures  $\mathbf{y}_r$  de dimension  $m_r$  linéaires en  $\mathbf{x}$  et entachées d'erreurs  $\mathbf{w}_r$  :

$$\mathbf{y}_r = \mathbf{H}_r \mathbf{x} + \mathbf{w}_r, r = 1 \text{ à } k$$

### 5.1 Estimation par moindres carrés classiques

Pour estimer  $\mathbf{x}$  on constitue un système linéaire global regroupant  $k$  lots de mesures  $\mathbf{y}_r$ , pour  $r = 1$  à  $k$ , sous la forme :

$$\mathbf{Y}_k = \mathbf{A}_k \mathbf{x} + \mathbf{W}_k$$

avec :

$$\mathbf{Y}_k = \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_k \end{pmatrix} \quad \mathbf{A}_k = \begin{pmatrix} \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_k \end{pmatrix} \quad \mathbf{W}_k = \begin{pmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_k \end{pmatrix}$$

$$\dim(\mathbf{Y}_k) = \dim(\mathbf{W}_k) = m = \sum_{r=1}^k m_r \text{ et } \dim(\mathbf{A}_k) = m \times n$$

Nous supposons que le nombre total de mesures  $m$  de mesures scalaires est supérieur au nombre d'inconnues  $n$  et que  $m$  est assez grand avec suffisamment de lignes indépendantes dans les  $\mathbf{H}_i$  pour que  $\text{rang}(\mathbf{A}_k) = n$ .

Recherchons  $\hat{\mathbf{x}}_k$  qui minimise la forme quadratique suivante :

$$C_k = \frac{1}{2} (\mathbf{Y}_k - \mathbf{A}_k \hat{\mathbf{x}}_k)^T \mathbf{R}_k^{-1} (\mathbf{Y}_k - \mathbf{A}_k \hat{\mathbf{x}}_k)$$

avec  $\mathbf{R}_k^{-1}$  matrice de pondération symétrique définie positive. La solution des moindres carrés (cf. chapitre précédent) s'écrit :

$$\hat{\mathbf{x}}_k = (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{Y}_k$$

En posant :

$$\begin{aligned} \mathbf{P}_k^{-1} &= \mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k \\ \mathbf{B}_k &= \mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{Y}_k \end{aligned}$$

le vecteur  $\hat{\mathbf{x}}_k$  solution du système linéaire symétrique :

$$\mathbf{P}_k^{-1} \hat{\mathbf{x}}_k = \mathbf{B}_k$$

s'écrit :

$$\hat{\mathbf{x}}_k = \mathbf{P}_k \mathbf{B}_k$$

## 5.2 Première forme récursive

Considérons un lot de mesures supplémentaires :

$$\mathbf{y}_{k+1} = \mathbf{H}_{k+1} \mathbf{x} + \mathbf{w}_{k+1}$$

et les nouveaux vecteurs et matrices globaux :

$$\mathbf{Y}_{k+1} = \begin{pmatrix} \mathbf{Y}_k \\ \mathbf{y}_{k+1} \end{pmatrix} \quad \mathbf{A}_{k+1} = \begin{pmatrix} \mathbf{A}_k \\ \mathbf{H}_{k+1} \end{pmatrix} \quad \mathbf{W}_{k+1} = \begin{pmatrix} \mathbf{W}_k \\ \mathbf{w}_{k+1} \end{pmatrix}$$

Considérons un nouveau critère sous la forme :

$$\begin{aligned} C_{k+1} &= C_k + \frac{1}{2} (\mathbf{y}_{k+1} - \mathbf{H}_{k+1} \hat{\mathbf{x}})^T \mathbf{r}_{k+1}^{-1} (\mathbf{y}_{k+1} - \mathbf{H}_{k+1} \hat{\mathbf{x}}) \\ &= \frac{1}{2} (\mathbf{Y}_{k+1} - \mathbf{A}_{k+1} \hat{\mathbf{x}})^T \mathbf{R}_{k+1}^{-1} (\mathbf{Y}_{k+1} - \mathbf{A}_{k+1} \hat{\mathbf{x}}) \end{aligned}$$

ce qui signifie que les matrices de pondération ont une forme bloc diagonale :

$$\mathbf{R}_{k+1}^{-1} = \begin{pmatrix} \mathbf{r}_1^{-1} & 0 & \dots & 0 \\ 0 & \mathbf{r}_2^{-1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{r}_{k+1}^{-1} \end{pmatrix} = \begin{pmatrix} \mathbf{R}_k^{-1} & 0 \\ 0 & \mathbf{r}_{k+1}^{-1} \end{pmatrix}$$

La nouvelle solution des moindres carrés s'écrit :

$$\begin{aligned} \hat{\mathbf{x}}_{k+1} &= (\mathbf{A}_{k+1}^T \mathbf{R}_{k+1}^{-1} \mathbf{A}_{k+1})^{-1} \mathbf{A}_{k+1}^T \mathbf{R}_{k+1}^{-1} \mathbf{Y}_{k+1} \\ &= \mathbf{P}_{k+1} \mathbf{B}_{k+1} \end{aligned}$$

c'est-à-dire sous la même forme que la précédente, avec :

$$\begin{aligned} \mathbf{P}_{k+1}^{-1} &= \mathbf{A}_{k+1}^T \mathbf{R}_{k+1}^{-1} \mathbf{A}_{k+1} = \sum_{r=1}^{k+1} \mathbf{H}_r^T \mathbf{r}_r^{-1} \mathbf{H}_r = \mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k + \mathbf{H}_{k+1}^T \mathbf{r}_{k+1}^{-1} \mathbf{H}_{k+1} \\ \mathbf{B}_{k+1} &= \mathbf{A}_{k+1}^T \mathbf{R}_{k+1}^{-1} \mathbf{Y}_{k+1} = \sum_{r=1}^{k+1} \mathbf{H}_r^T \mathbf{r}_r^{-1} \mathbf{y}_r = \mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{Y}_k + \mathbf{H}_{k+1}^T \mathbf{r}_{k+1}^{-1} \mathbf{y}_{k+1} \end{aligned}$$

d'où les relations de récurrence pour les des deux principaux facteurs  $\mathbf{P}_k^{-1}$  et  $\mathbf{B}_k$  :

$$\begin{aligned} \mathbf{P}_{k+1}^{-1} &= \mathbf{P}_k^{-1} + \mathbf{H}_{k+1}^T \mathbf{r}_{k+1}^{-1} \mathbf{H}_{k+1} \\ \mathbf{B}_{k+1} &= \mathbf{B}_k + \mathbf{H}_{k+1}^T \mathbf{r}_{k+1}^{-1} \mathbf{y}_{k+1} \end{aligned}$$

### 5.3 Deuxième forme récursive

C'est cette deuxième forme que l'on désigne habituellement par "*moindres carrés récursifs*".

Pour simplifier les écritures, on considère les étapes  $k$  et  $k - 1$ . Les relations de récurrence précédentes s'écrivent :

$$\begin{aligned}\mathbf{P}_k &= (\mathbf{P}_{k-1}^{-1} + \mathbf{H}_k^T \mathbf{r}_k^{-1} \mathbf{H}_k)^{-1} \\ \mathbf{B}_k &= \mathbf{B}_{k-1} + \mathbf{H}_k^T \mathbf{r}_k^{-1} \mathbf{y}_k\end{aligned}$$

En appliquant le deuxième lemme d'inversion matricielle à la première de ces deux relations, on obtient :

$$\mathbf{P}_k = \mathbf{P}_{k-1} - \mathbf{P}_{k-1} \mathbf{H}_k^T (\mathbf{r}_k + \mathbf{H}_k \mathbf{P}_{k-1} \mathbf{H}_k^T)^{-1} \mathbf{H}_k \mathbf{P}_{k-1} \quad (5.1)$$

Généralement, on pose :

$$\mathbf{K}_k = \mathbf{P}_{k-1} \mathbf{H}_k^T (\mathbf{r}_k + \mathbf{H}_k \mathbf{P}_{k-1} \mathbf{H}_k^T)^{-1} \quad (5.2)$$

qui s'écrit également (en appliquant le premier lemme d'inversion matricielle) :

$$\mathbf{K}_k = (\mathbf{P}_{k-1}^{-1} + \mathbf{H}_k^T \mathbf{r}_k^{-1} \mathbf{H}_k)^{-1} \mathbf{H}_k^T \mathbf{r}_k^{-1} \quad (5.3)$$

$$\mathbf{K}_k = \mathbf{P}_k \mathbf{H}_k^T \mathbf{r}_k^{-1} \quad (5.4)$$

En utilisant (5.2), la relation (5.1) se met sous la forme :

$$\mathbf{P}_k = \mathbf{P}_{k-1} - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{k-1} \quad (5.5)$$

La solution  $\hat{\mathbf{x}}_k$  s'écrit :

$$\begin{aligned}\hat{\mathbf{x}}_k &= \mathbf{P}_k \mathbf{B}_k = \mathbf{P}_k (\mathbf{B}_{k-1} + \mathbf{H}_k^T \mathbf{r}_k^{-1} \mathbf{y}_k) \\ &= \mathbf{P}_k \mathbf{B}_{k-1} + \mathbf{P}_k \mathbf{H}_k^T \mathbf{r}_k^{-1} \mathbf{y}_k\end{aligned}$$

En utilisant (5.5) pour le premier terme et (5.4) pour le deuxième, on obtient :

$$\begin{aligned}\hat{\mathbf{x}}_k &= (\mathbf{P}_{k-1} - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{k-1}) \mathbf{B}_{k-1} + \mathbf{K}_k \mathbf{y}_k \\ \hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_{k-1} - \mathbf{K}_k \mathbf{H}_k \hat{\mathbf{x}}_{k-1} + \mathbf{K}_k \mathbf{y}_k \\ &= \hat{\mathbf{x}}_{k-1} + \mathbf{K}_k (\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k-1})\end{aligned}$$

En résumé, la deuxième forme récursive est (traditionnellement) constituée des équations suivantes :

$$\left\{ \begin{array}{l} \mathbf{K}_k = \mathbf{P}_{k-1} \mathbf{H}_k^T (\mathbf{r}_k + \mathbf{H}_k \mathbf{P}_{k-1} \mathbf{H}_k^T)^{-1} \\ \hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k-1} + \mathbf{K}_k (\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k-1}) \\ \mathbf{P}_k = \mathbf{P}_{k-1} - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{k-1} \end{array} \right. \quad (5.6)$$

Pour initialiser cette procédure récursive, il faut disposer d'un premier lot de mesures suffisant pour appliquer la forme non récursive, à savoir :

$$\begin{aligned}\mathbf{P}_k &= (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \\ \hat{\mathbf{x}}_k &= \mathbf{P}_k \mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{Y}_k\end{aligned}$$

## 5.4 Comparaison des deux formes récursives

La première forme récursive met en oeuvre des formules plus simples, mais nécessite l'inversion à chaque itération de la matrice  $\mathbf{P}_k$  de dimension  $n \times n$ .

La deuxième forme, met en oeuvre des formules un peu plus complexes, mais ne nécessite à chaque itération que l'inversion d'une matrice  $m_k \times m_k$  qui se réduit à un simple scalaire si on traite les mesures composante par composante, et c'est là l'avantage majeur des moindres carrés récursifs.

## 5.5 Liaison avec les formules du filtre de Kalman discret

Les équations du filtre de Kalman discret sont exactement les mêmes que celles des moindres carrés récursifs (5.6). Le filtre de Kalman discret permet d'estimer un vecteur  $\mathbf{x}$  à partir de mesures :

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{w}_k$$

où  $\mathbf{w}_k$  est un bruit de mesure pseudo-blanc gaussien de moyenne nulle, sans corrélation entre lots de mesures. Les matrices  $\mathbf{r}_k$  qui interviennent dans les équations du filtre de Kalman sont les matrices de covariance du bruit de mesure, telles que :

$$E(\mathbf{w}_k \mathbf{w}_j^T) = \mathbf{r}_k \delta_{kj} \quad \forall k, j$$

Dans le cas du filtre de Kalman, la matrice  $\mathbf{P}_k$  représente la matrice de covariance de l'erreur d'estimation, à savoir :

$$\mathbf{P}_k = E\left((\hat{\mathbf{x}}_k - \mathbf{x})^T (\hat{\mathbf{x}}_k - \mathbf{x})\right)$$

Dans le cas des moindres carrés récursifs, la matrice  $\mathbf{r}_k^{-1}$  n'est qu'une matrice de pondération symétrique positive laissée au choix de l'utilisateur et la matrice  $\mathbf{P}_k$  est un intermédiaire de calcul. Examinons leurs significations pratiques a posteriori.

### 5.5.1 Covariance de l'erreur d'estimation dans le cas de moindres carrés récursifs

Dans le cas des moindres carrés récursifs,  $\mathbf{w}_k$  est une erreur de mesure quelconque. Toutefois pour obtenir une bonne estimation  $\hat{\mathbf{x}}_k$  il y a intérêt à donner aux pondérations  $\mathbf{r}_k$  des valeurs qui sont en rapport avec la précision des mesures  $\mathbf{y}_k$ . Donnons à cette matrice la même signification que dans le cas filtre de Kalman, à savoir, la covariance de l'erreur de mesure :

$$E(\mathbf{w}_k \mathbf{w}_k^T) = \mathbf{r}_k$$

éventuellement non diagonale (symétrique définie positive). Nous supposons donc que les erreurs de mesure sont à moyenne nulle :

$$E(\mathbf{w}_k) = 0$$

Avec ces conditions, calculons l'erreur d'estimation  $\Delta \mathbf{x}_k = \hat{\mathbf{x}}_k - \mathbf{x}$  dans le cas de la méthode des moindres carrés. Il vient :

$$\begin{aligned} \Delta \mathbf{x}_k &= (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{Y}_k - \mathbf{x} \\ &= (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \mathbf{A}_k^T \mathbf{R}_k^{-1} (\mathbf{A}_k \mathbf{x} + \mathbf{W}_k) - \mathbf{x} \\ &= (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{W}_k \end{aligned}$$

Calculons l'erreur moyenne d'estimation :

$$E(\Delta \mathbf{x}_k) = E \left[ (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{W}_k \right] = (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \mathbf{A}_k^T \mathbf{R}_k^{-1} E(\mathbf{W}_k) = 0$$

Si l'erreur de mesure est à moyenne nulle, l'erreur d'estimation l'est également.

Calculons la covariance de l'erreur d'estimation :

$$\begin{aligned} E(\Delta \mathbf{x}_k \Delta \mathbf{x}_k^T) &= E \left[ (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{W}_k \mathbf{W}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \right] \\ &= (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \mathbf{A}_k^T \mathbf{R}_k^{-1} E(\mathbf{W}_k \mathbf{W}_k^T) \mathbf{R}_k^{-1} \mathbf{A}_k (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \\ &= (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{R}_k \mathbf{R}_k^{-1} \mathbf{A}_k (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \\ &= (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \\ &= \mathbf{P}_k \end{aligned}$$

La matrice  $\mathbf{P}_k$  des moindres carrés récursifs est (comme dans le filtre de Kalman) la matrice de covariance de l'erreur d'estimation, sous les seules hypothèses que les erreurs de mesure sont à moyenne nulle et qu'il n'y a pas de corrélation entre les erreurs des différents lots. L'hypothèse gaussienne sur les bruits de mesure, nécessaire dans le cas du filtre de Kalman, n'est pas nécessaire ici.

### 5.5.2 Estimation de la variance des erreurs de mesure

Si toutes les mesures sont de même nature, avec une erreur à moyenne nulle et avec le même écart type  $\sigma$ , cet écart-type peut être estimé a posteriori.

Considérons l'erreur entre mesure réelle et mesure estimée :

$$\begin{aligned} \Delta \mathbf{Y}_k &= \mathbf{Y}_k - \mathbf{A}_k \hat{\mathbf{x}}_k \\ &= \mathbf{A}_k \mathbf{x}_k + \mathbf{W}_k - \mathbf{A}_k \hat{\mathbf{x}}_k \\ &= \mathbf{A}_k \mathbf{x}_k + \mathbf{W}_k - \mathbf{A}_k (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{Y}_k \\ &= \mathbf{A}_k \mathbf{x}_k + \mathbf{W}_k - \mathbf{A}_k (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \mathbf{A}_k^T \mathbf{R}_k^{-1} (\mathbf{A}_k \mathbf{x}_k + \mathbf{W}_k) \\ &= \left( \mathbf{I} - \mathbf{A}_k (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \mathbf{A}_k^T \mathbf{R}_k^{-1} \right) \mathbf{W}_k \\ &= \mathbf{G}_k \mathbf{W}_k \end{aligned}$$

où on a introduit la matrice symétrique :

$$\mathbf{G}_k = \mathbf{I} - \mathbf{A}_k (\mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k)^{-1} \mathbf{A}_k^T \mathbf{R}_k^{-1}$$

En constatant que :

$$\mathbf{G}_k^2 = \mathbf{G}_k$$

Il vient :

$$E(\|\Delta \mathbf{Y}_k\|^2) = E(\Delta \mathbf{Y}_k^T \Delta \mathbf{Y}_k) = E(\mathbf{W}_k^T \mathbf{G}_k^2 \mathbf{W}_k) = E(\mathbf{W}_k^T \mathbf{G}_k \mathbf{W}_k)$$

Dans le cas particuliers où tous les  $w_i$  sont indépendants (indépendance entre lots et entre composantes), on a :

$$E(\|\Delta \mathbf{Y}_k\|^2) = \sum_{i=1}^m g_{ii} E(w_i^2)$$

De plus si  $E(w_i^2) = cte = \sigma^2$  alors :

$$E\left(\|\Delta \mathbf{Y}_k\|^2\right) = \sigma^2 \sum_{i=1}^m g_{ii} = \sigma^2 \left[ m - \text{trace} \left( \mathbf{A}_k \left( \mathbf{A}_k^T \mathbf{R}_k^{-1} \mathbf{A}_k \right)^{-1} \mathbf{A}_k^T \mathbf{R}_k^{-1} \right) \right]$$

avec par hypothèse  $\mathbf{R}_k^{-1} = \frac{1}{\sigma^2} \mathbf{I}$ . D'où :

$$\begin{aligned} E\left(\|\Delta \mathbf{Y}_k\|^2\right) &= \sigma^2 \left[ m - \text{trace} \left( \mathbf{A}_k \left( \mathbf{A}_k^T \frac{1}{\sigma^2} \mathbf{A}_k \right)^{-1} \mathbf{A}_k^T \frac{1}{\sigma^2} \right) \right] \\ &= \sigma^2 \left[ m - \text{trace} \left( \mathbf{A}_k \left( \mathbf{A}_k^T \mathbf{A}_k \right)^{-1} \mathbf{A}_k^T \right) \right] \\ &= (m - n) \sigma^2 \end{aligned}$$

Ainsi, si toutes les mesures sont de même nature, on peut estimer la variance  $\sigma^2$  de l'erreur de mesure à partir de cette équation :

$$\hat{\sigma}^2 = \frac{1}{m - n} \sum_{i=1}^m (y_i - \hat{y}_i)^2$$

## 5.6 Compléments sur l'estimation

Pour compléter la liaison entre moindres carrés récursif et filtre de Kalman, nous présentons les formules utilisées pour le processus de propagation dans le cas des systèmes linéaires discrets.

### 5.6.1 La propagation de l'estimation

Le filtre de Kalman discret classique fait généralement intervenir un processus dynamique qui fait évoluer l'état dans les intervalles qui séparent les lots mesures  $\mathbf{y}_{k-1}$  et  $\mathbf{y}_k$ . Considérons le modèle linéaire discret suivant :

$$\mathbf{x}_k = \mathbf{F}_{k-1} \mathbf{x}_{k-1} + \mathbf{u}_{k-1} + \mathbf{v}_{k-1}$$

où  $\mathbf{u}_k$  est un vecteur connu et où  $\mathbf{v}_k$  est une erreur de modélisation à moyenne nulle et de covariance connue  $\mathbf{Q}_k$  telle que :

$$E(\mathbf{v}_k \mathbf{v}_j^T) = \mathbf{Q}_k \delta_{ki}$$

c'est-à-dire sans corrélation entre les différentes étapes.

Le modèle de ce processus permet de faire une nouvelle estimation  $\hat{\mathbf{x}}_{k/k-1}$  (ce qui signifie estimée de  $\mathbf{x}_k$  à partir des informations disponibles à l'étape  $k-1$ ), avant la correction prenant en compte la mesure  $\mathbf{y}_k$  :

$$\hat{\mathbf{x}}_{k/k-1} = \mathbf{F}_{k-1} \hat{\mathbf{x}}_{k-1} + \mathbf{u}_{k-1}$$

Ce calcul est appelé propagation de l'estimation.

Notons  $\Delta_{k-1}$  l'erreur d'estimation avant cette phase de propagation, et après la correction qui a utilisé la mesure  $\mathbf{y}_{k-1}$  :

$$\Delta_{k-1} = \mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}$$

et notons  $\Delta_{k/k-1}$  l'erreur d'estimation à l'issue de cette phase de propagation :

$$\Delta_{k/k-1} = \mathbf{x}_k - \hat{\mathbf{x}}_{k/k-1} = \mathbf{F}_{k-1} \Delta_{k-1} + \mathbf{v}_{k-1}$$

Par définition, on a :

$$E(\Delta_{k-1}) = 0 \text{ et } E(\Delta_{k-1} \Delta_{k-1}^T) = \mathbf{P}_{k-1}$$



Evaluons  $E(\Delta_{k/k-1})$  et  $\mathbf{P}_{k/k-1} = E(\Delta_{k/k-1}\Delta_{k/k-1}^T)$ . Il vient :

$$E(\Delta_{k/k-1}) = \mathbf{F}_{k-1}E(\Delta_{k-1}) + E(\mathbf{v}_{k-1}) = 0$$

et :

$$\begin{aligned} E(\Delta_{k/k-1}\Delta_{k/k-1}^T) &= \mathbf{F}_{k-1}E(\Delta_{k-1}\Delta_{k-1}^T)\mathbf{F}_{k-1}^T + E(\mathbf{v}_k\mathbf{v}_k^T) \\ &\quad + E(\mathbf{v}_{k-1}\Delta_{k-1}^T)\mathbf{F}_{k-1}^T + \mathbf{F}_{k-1}E(\Delta_{k-1}\mathbf{v}_{k-1}^T) \end{aligned}$$

Les termes de la deuxième ligne sont nuls car on a considéré qu'il n'y a pas de corrélation entre les différentes étapes. Il vient :

$$\mathbf{P}_{k/k-1} = \mathbf{F}_{k-1}\mathbf{P}_{k-1}\mathbf{F}_{k-1}^T + \mathbf{Q}_{k-1}$$

qui est la formule de propagation de la covariance.

Lorsque  $\mathbf{y}_k$  est pris en compte, les formules correctives sont appliquées aux éléments prédits par cette phase de propagation :

$$\begin{aligned} \mathbf{K}_k &= \mathbf{P}_{k/k-1}\mathbf{H}_k^T(\mathbf{r}_k + \mathbf{H}_k\mathbf{P}_{k/k-1}\mathbf{H}_k^T)^{-1} \\ \hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_{k/k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k/k-1}) \\ \mathbf{P}_k &= \mathbf{P}_{k/k-1} - \mathbf{K}_k\mathbf{H}_k\mathbf{P}_{k/k-1} \end{aligned}$$

## 5.6.2 La formule de Joseph

Considérons la formule de correction de la covariance d'erreur d'estimation (ré-écrite en notation simplifiée sans phase de propagation). Si  $\mathbf{P}_{k-1}$  est symétrique définie positive, la formule :

$$\mathbf{P}_k = \mathbf{P}_{k-1} - \mathbf{K}_k\mathbf{H}_k\mathbf{P}_{k-1}$$

ne met pas en évidence le fait que  $\mathbf{P}_k$  l'est également. Calculons :

$$\begin{aligned} \mathbf{K}_k\mathbf{H}_k\mathbf{P}_{k-1} &= \mathbf{K}_k(\mathbf{r}_k + \mathbf{H}_k\mathbf{P}_{k-1}\mathbf{H}_k^T)(\mathbf{r}_k + \mathbf{H}_k\mathbf{P}_{k-1}\mathbf{H}_k^T)^{-1}\mathbf{H}_k\mathbf{P}_{k-1} \\ &= \mathbf{K}_k(\mathbf{r}_k + \mathbf{H}_k\mathbf{P}_{k-1}\mathbf{H}_k^T)\mathbf{K}_k^T \\ &= \mathbf{P}_{k-1}\mathbf{H}_k^T\mathbf{K}_k^T \end{aligned}$$

D'où :

$$\begin{aligned} \mathbf{P}_k &= \mathbf{P}_{k-1} - \mathbf{K}_k\mathbf{H}_k\mathbf{P}_{k-1} = \mathbf{P}_{k-1} - \mathbf{K}_k\mathbf{H}_k\mathbf{P}_{k-1} + \mathbf{K}_k(\mathbf{r}_k + \mathbf{H}_k\mathbf{P}_{k-1}\mathbf{H}_k^T)\mathbf{K}_k^T - \mathbf{P}_{k-1}\mathbf{H}_k^T\mathbf{K}_k^T \\ \mathbf{P}_k &= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k-1}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^T + \mathbf{K}_k\mathbf{r}_k\mathbf{K}_k^T \end{aligned}$$

Cette formule de mise à jour (dite *formule de Joseph*) est mieux conditionnée pour conserver la symétrie et la positivité. Toutefois elle est beaucoup plus complexe. Pour les processus bien conditionnés, il est plus simple de symétriser la matrice  $\mathbf{P}_k$  en la remplaçant par  $\frac{1}{2}(\mathbf{P}_k + \mathbf{P}_k^T)$ . Dans le cas de systèmes très mal conditionnés, il existe des techniques plus fiables comme celle de Bierman et ThornTon qui utilisent des formules récursives avec une décomposition  $\mathbf{UDU}^T$  de la matrice  $\mathbf{P}$ .

### 5.6.3 Gain de Kalman en continu et en discret

Le gain de Kalman  $\mathbf{K}_k$  en discret est donné en fonction de  $\mathbf{P}_{k-1}$  par la formules (5.2) et en fonction de  $\mathbf{P}_k$  par la formule (5.4), à savoir :

$$\mathbf{K}_k = \mathbf{P}_k \mathbf{H}_k^T \mathbf{r}_k^{-1}$$

Cette dernière expression s'apparente à celle du gain de Kalman  $\mathbf{K}(t)$  pour les processus continus. Elle utilise la matrice de covariance  $\mathbf{P}_{\hat{x}}$  qui prend en compte la correction effectuée en continu. Mais la comparaison ne doit pas être poussée plus loin, car dans le cas discret,  $\mathbf{K}_k(\mathbf{y} - \mathbf{A}\mathbf{x}_{k-1})$  corrige  $\mathbf{x}_{k-1}$  alors que dans le filtre de Kalman continu  $\mathbf{K}(\mathbf{y} - \mathbf{A}\mathbf{x})$  corrige la dérivée  $\dot{\mathbf{x}}$  de  $\mathbf{x}$ . Les deux gains ne sont donc pas comparables.

## Chapitre 6

# Les bruits dans les filtres de Kalman

La liaison établie au chapitre précédent entre moindres carrés récursifs et filtre de Kalman, nous montre qu'on peut établir un filtre de Kalman en choisissant arbitrairement les matrices  $\mathbf{r}_k$ . Toutefois ces matrices (et les matrices  $\mathbf{Q}_k$  s'il y a un processus de propagation) conditionnent fortement la qualité du résultat. Il en résulte que l'interprétation physique de ces covariances est un guide précieux pour le réglage du filtre. Les utilisateurs débutants sont toutefois souvent déroutés par les unités utilisées pour exprimer ces covariances dans les notices des appareils de mesure. Ce chapitre a pour objet de préciser la signification de ces unités.

### 6.1 Le bruit d'état

#### 6.1.1 Système continu

Soit un système réel modélisé par l'équation d'état *scalaire* :

$$\dot{x}(t) = e(t) + v(t)$$

et son estimateur modélisé par :

$$\dot{\hat{x}}(t) = e(t)$$

où  $e(t)$  est une entrée déterministe connue à moyenne nulle et  $v(t)$  un *bruit blanc gaussien stationnaire à moyenne nulle* dont la fonction d'autocorrélation est supposée être modélisée par :

$$E[v(t)v(t+\tau)] = R(\tau) = Q_c \delta(\tau) \text{ avec } Q_c = \Phi_{vv}(\omega) = cte$$

où l'indice  $c$  rappelle le cas Continu,  $R(\tau)$  est la fonction d'autocorrélation de  $v(t)$  et  $\Phi_{vv}(\omega)$  est sa densité spectrale de puissance (ou plus simplement son spectre de puissance), transformée de Fourier de la fonction d'autocorrélation.

Notons  $U$  l'unité utilisée pour mesurer  $x$ . Alors :

- $v$  s'exprime en  $UT^{-1}$
- $Q_c \delta(\tau)$  en  $U^2 T^{-2}$ .
- $\Phi$  et  $Q_c$  s'exprime en  $U^2 T^{-1}$  (car  $Q_c = \int_{-\infty}^{+\infty} Q_c \delta(\tau) d\tau$ ).

La formule de Parseval :

$$R(0) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} x^2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Phi(\omega) d\omega = P_m$$

montre que la puissance moyenne d'un bruit blanc est infinie (ainsi que son énergie). En pratique, les bandes passantes étant limitées, la puissance du bruit réel est bornée. Généralement le bruit est caractérisé par  $\sqrt{Q_c}$ , que nous appellerons pseudo écart-type de  $v(t)$ .

Posons :

$$\varepsilon = x(t) - \hat{x}(t) \rightarrow \dot{\varepsilon}(t) = v(t)$$

et notons :

$$\Lambda_c(t) = E[\varepsilon^2(t)]$$

la valeur moyenne de la puissance de  $\varepsilon(t)$  qui s'exprime en  $U^2$ .

Si  $e(t)$ , puis  $\varepsilon(t)$  sont à moyenne nulle, on démontre<sup>1</sup> que :

$$\dot{\Lambda}_c(t) = Q_c \rightarrow \Lambda_c(t) = Q_c t$$

### 6.1.2 Système discret

Une représentation numérique de ce processus sera effectuée, par exemple, par le modèle suivant :

$$\varepsilon_n = \varepsilon_{n-1} + W_{n-1}$$

où  $W$  a pour dimension  $U$ , comme  $\varepsilon$  et  $x$ .

Pour cette représentation, on suppose que  $e(t)$  est échantillonné à la période  $T$  et que la valeur échantillonnée est bloquée pendant toute cette période.

Supposons que :

$$E[W_n W_k] = Q_d \delta_{nk}$$

où l'indice  $d$  rappelle le cas discret. Ici la dimension de  $Q_d$  est simplement celle de  $W^2$  c'est-à-dire  $U^2$  :

—  $Q_d$  s'exprime en  $U^2$ .

Posons :

$$E[\varepsilon_n^2] = \Lambda_{dn}$$

de dimension  $U^2$  et supposons que  $E[\varepsilon_n \varepsilon_k] = \Lambda_{dn} \delta_{nk}$ .

Si les  $\varepsilon_n$  sont à moyenne nulle, on démontre que :

$$\Lambda_{dn} = \Lambda_{dn-1} + Q_d \rightarrow \Lambda_{dn} = nQ_d$$

(tout est de dimension  $U^2$ ).

Si on désire que  $\Lambda_{dn}$  représente la covariance de  $\varepsilon(t)$  à l'instant  $nT$ , c'est-à-dire si on veut que  $\Lambda_{dn} = \Lambda_c(nT)$ , il faut que  $nQ_d = Q_c nT$ , c'est-à-dire que :

$$Q_d = Q_c T$$

En conclusion, si  $x$  s'exprime en  $U$ , la covariance du bruit d'état en discret s'exprime en  $U^2$  et la covariance du bruit d'état en continu s'exprime en  $U^2 T^{-1}$ . Si on discrétise un processus continu de covariance  $Q_c$  avec une période d'échantillonnage  $T$ , la covariance du processus discret peut être approximée par  $Q_d = Q_c T$ .

1. voir justification approximative en annexe

### 6.1.3 Simulation numérique approchée du processus continu

Pour simuler numériquement l'intégration du processus continu  $\dot{\varepsilon}(t) = v(t)$ , on a en première approximation :

$$\varepsilon(t+T) = \varepsilon(t) + v'(t)T$$

où  $v'(t)$  est un signal qui est bloqué pendant la durée  $T$ , alors que  $v(t)$  varie pendant ces périodes de blocage. Nous utilisons l'exposant  $'$  pour distinguer  $v'(t)$  qui est une fonction en escalier de la fonction continue  $v(t)$ .  $v'(t)$  et  $v(t)$  coïncident aux instants d'échantillonnage. Remarquons au passage que tous les signaux de la simulation numérique approchée ( $e(t)$ ,  $\varepsilon(t)$ ,  $x(t)$ , ...) sont des signaux en escalier

En comparant avec le processus purement discret, il vient :

$$W_n = v'(nT)T$$

$$E[W_n W_k] = E[v'(nT)T v'(kT)T] = T^2 E[v'(nT)v'(kT)]$$

Or :

$$W_n = v'(nT)T$$

$$E[W_n W_k] = Q_d \delta_{nk}$$

d'où :

$$E[v'^2(nT)] = \frac{1}{T^2} Q_d = \frac{1}{T} Q_c$$

Notons  $R'(\tau) = E[v'(t)v'(t+\tau)]$  la fonction d'autocorrélation du signal en escalier  $v'(t)$ . C'est un triangle de base  $[-T, T]$  et d'amplitude  $E[v'^2(t)] = \frac{1}{T} Q_c$ .

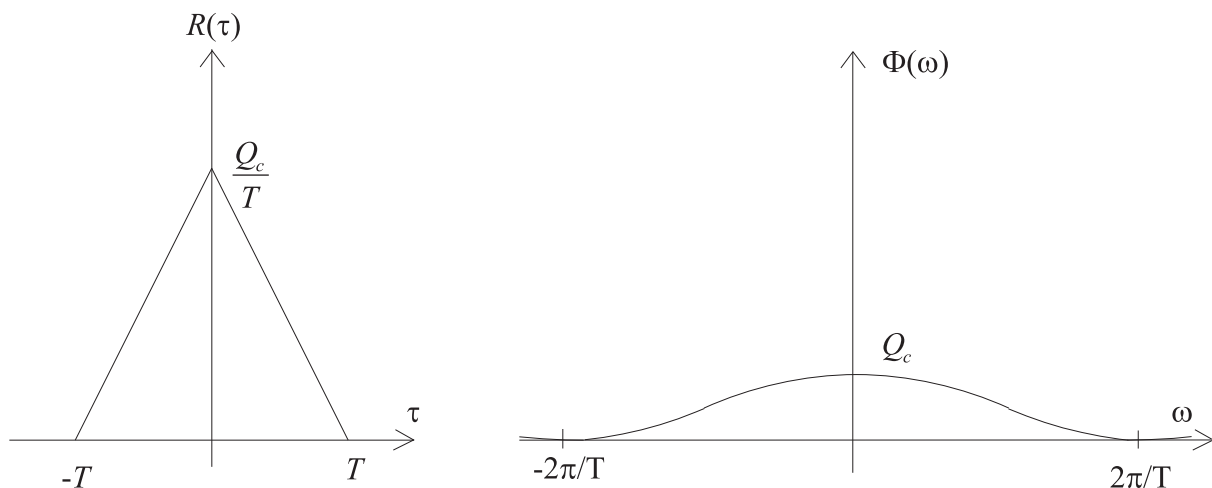


FIGURE 6.1 – Fonction d'autocorrélation et spectre de puissance de  $v'(t)$

La densité spectrale de puissance associée s'écrit :

$$\Phi_{v'v'}(\omega) = T \frac{\sin\left(\frac{\omega T}{2}\right)}{\left(\frac{\omega T}{2}\right)^2} \cdot \frac{1}{T} Q_c = Q_c \frac{\sin^2\left(\frac{\omega T}{2}\right)}{\left(\frac{\omega T}{2}\right)^2}$$

Le signal  $v'(t)$  a donc une énergie qui décroît de  $Q_c$  à 0 dans la bande de fréquence de 0 à  $f = 1/T$ .

### 6.1.4 Conclusion sur les représentations de bruit d'état

Il y a 3 représentation du bruit d'état :

1.  $v(t)$  le bruit blanc du système continu, de variance infini (et d'énergie infinie) , mais de densité spectrale constante  $Q_c$  qui affecte  $\dot{\varepsilon}(t)$ .  $\sqrt{Q_c}$ , son pseudo écart-type peut s'exprimer soit en  $U\sqrt{T^{-1}}$  c'est-à-dire en  $U/\sqrt{T}$
2.  $v'(t)$  le bruit blanc échantillonné, bloqué pendant la période  $T$ , de variance  $\frac{1}{T}Q_c = \frac{1}{T^2}Q_d$ , de même puissance moyenne  $Q_c$  mais dans une bande de fréquence limitée par l'échantillonnage (donc d'énergie finie). Ce bruit  $v'(t)$  est ajouté à  $\varepsilon(t)$  après avoir été multiplié par  $T$ . Son écart-type s'exprime en  $UT^{-1}$ .
3.  $W_n$  le bruit blanc numérique de variance  $Q_d = Q_cT$ , qui est *directement* ajouté à  $x_n$ . Son écart-type est de dimension  $U$  (comme  $\varepsilon$  auquel  $W$  est directement ajouté).

## 6.2 Le bruit de mesure

Supposons qu'une équation de mesure soit modélisée par :

$$z(t) = x(t) + w(t)$$

en continu et par :

$$z_n = x_n + w_n$$

en discret ou  $w(t)$  est un bruit de mesure blanc gaussien et  $w_n$  un bruit numérique pseudo-blanc tels que :

$$\begin{aligned} E[w(t)w(t+\tau)] &= R_c\delta(\tau) \text{ avec } R_c = \Phi_{ww}(\omega) = cte \\ E[w_n w_k] &= R_d\delta_{nk} \end{aligned}$$

Dans le filtre de Kalman continu, la mesure est utilisée pour corriger  $\hat{x}(t)$  par l'équation :

$$\dot{\hat{x}}(t) = e(t) + K_c(z(t) - \hat{x}(t))$$

avec :

$$K_c = \frac{\Lambda_c}{R_c}$$

Une simulation numérique simplifiée de ce filtre s'écrirait :

$$\hat{x}(t+T) = \hat{x}(t) + e(t)T + K_c(z'(t) - \hat{x}(t))T \quad (6.1)$$

alors que l'implantation du filtre de Kalman *directement* en numérique s'écrit :

$$\hat{x}_n = \hat{x}_{n-1} + e_{n-1}T + K_{dn}[z_n - (\hat{x}_{n-1} + e_{n-1}T)] \quad (6.2)$$

avec :

$$K_{dn} = \frac{\Lambda_{dn}^-}{\Lambda_{dn}^- + R_d} = \frac{\Lambda_{dn}^+}{R_d}$$

où les covariances  $\Lambda_{dn}^+$  après et  $\Lambda_{dn}^-$  avant correction sont liées par :

$$\Lambda_{dn}^+ = \Lambda_{dn}^- - K_{dn}\Lambda_{dn}^- = \frac{\Lambda_{dn}^- R_d}{\Lambda_{dn}^- + R_d} = K_{dn}R_d$$

En comparant la réalisation discrète (6.2) à la réalisation continue simulée en numérique (6.1), on voit que les gains des filtres continu et discret sont liés par la relation :

$$K_{dn} = K_c T$$

Si on veut que  $\Lambda_{dn}^+$  après correction soit de l'ordre de  $\Lambda_c$  :

$$\begin{aligned}\Lambda_{dn}^+ &= K_{dn} R_d = (K_c T) R_d \\ \Lambda_c &= K_c R_c\end{aligned}$$

on doit avoir :

$$R_d \simeq \frac{1}{T} R_c$$

La variance du bruit utilisé dans la simulation numérique va tendre vers l'infini quand  $T$  tendra vers zéro.

**Remarque :** La formule

$$\frac{1}{\Lambda_{dn}^+} = \frac{1}{\Lambda_{dn}^-} + \frac{1}{R_d}$$

montre que :

$$\Lambda_{dn}^+ \leq \min(\Lambda_{dn}^-, R_d)$$

La qualité du résultat après correction est au moins égale à la meilleure des qualités.

### 6.2.1 Conclusion sur les représentations de bruit de mesure

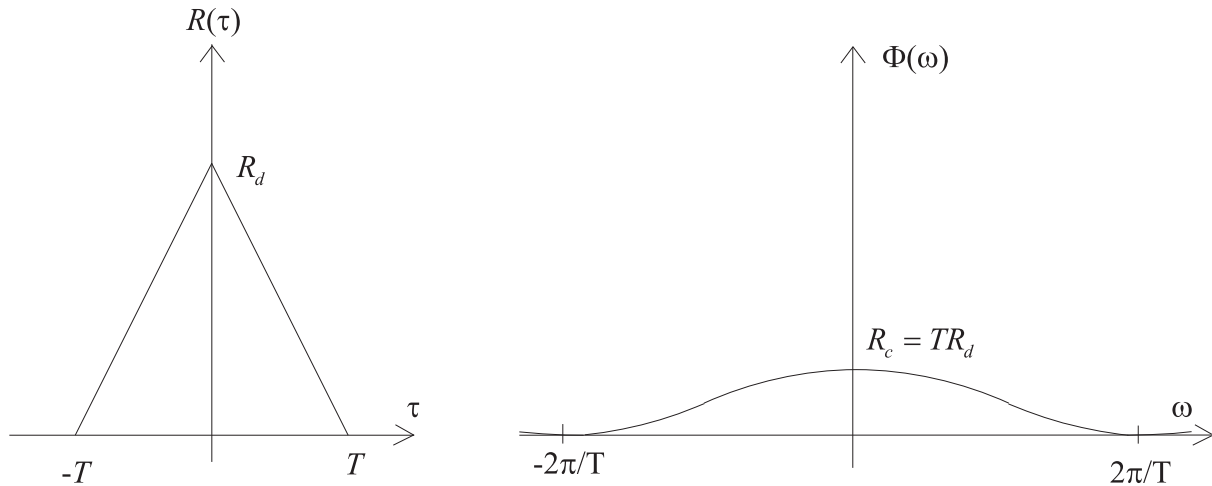
Ici aussi, il y a 3 représentation du bruit de mesure :

1.  $w_n$  le bruit blanc numérique de variance  $R_d$ , qui s'ajoute *directement* à  $z_n$ . Son écart-type est de dimension U (comme  $z$  auquel  $w$  est directement ajouté) et sa covariance  $R_d$  s'exprime en  $U^2$ .
2.  $w(t)$  le bruit blanc qui s'ajoute à la mesure effectuée en continu, de variance infinie (et d'énergie infinie), mais de densité spectrale constante  $R_c = T R_d$ . Son pseudo écart-type  $\sqrt{R_c}$  peut s'exprimer en  $U\sqrt{T}$ , ou en  $U/\sqrt{T^{-1}}$  que l'on trouve sous la forme  $U/\sqrt{\text{Hz}}$  dans les notices des constructeurs.
3.  $w'(t)$  le bruit blanc continu échantillonné, bloqué pendant la période  $T$ , de même variance  $R_d$  que  $w_n$ . Son écart-type s'exprime en U comme la mesure  $z$ . Sa fonction d'autocorrélation est un triangle de base  $[-T, +T]$  et d'amplitude  $R_d$  (son aire est ainsi bien égale à  $R_c$  qui est l'aire sous la fonction d'autocorrélation de  $w(t)$  :  $R(\tau) = R_c \delta(\tau)$ ). Son spectre a une amplitude voisine de  $R_c = T R_d$  dans la bande de fréquence  $[-\frac{1}{2T}, \frac{1}{2T}]$ . Il s'annule pour  $f = \pm \frac{1}{T}$  et tend vers 0 à l'infini. Sa puissance moyenne dans la bande de fréquence  $[-\frac{1}{2T}, \frac{1}{2T}]$  vaut  $\frac{1}{T} (T R_d) = R_d$  voisine celle du bruit blanc qui dans la même bande vaut  $\frac{1}{T} R_c$ . Les énergies des deux signaux sont bien les mêmes dans la bande  $[-\frac{1}{2T}, \frac{1}{2T}]$ .

Remarque : Lorsque le bruit de mesure  $w(t)$  est tel que sa densité spectrale de puissance est limité dans une bande passante  $f_c$ , avec  $\Phi_{ww}(\omega)$  relativement constant et égal à  $R_c$  dans une bonne partie de cette bande passante, sa puissance moyenne est approximativement donnée par :

$$P_m = R(0) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Phi(\omega) d\omega \simeq f_c R_c$$

Cette relation montre que  $\sqrt{R_c} \simeq \sqrt{P_m}/\sqrt{f_c}$ . Dans cette relation  $\sqrt{P_m}$  s'exprime en unités U de la variable mesurée  $z$  et  $\sqrt{f_c}$  en  $\sqrt{\text{Hz}}$ . Le niveau du bruit caractérisé par le pseudo écart-type  $\sqrt{R_c}$  s'exprime ainsi en unités U de la variable mesurée par  $\sqrt{\text{Hz}}$ .

FIGURE 6.2 – Autocorrélation et spectre de puissance de  $w'(t)$ 

### 6.3 Exemple : le gyromètre

Considérons à titre d'exemple le cas d'une mesure gyrométrique  $\omega_m$  modélisée par :

$$\omega_m(t) = \omega(t) + b_\omega(t)$$

où  $b_\omega(t)$  est par exemple un bruit blanc d'origine électronique, et où  $\omega$  est exprimé en  $U = \text{deg/s}$ . Le bruit  $b_\omega(t)$  va être caractérisé par son pseudo écart-type  $\sigma_\omega = \sqrt{R_c}$  exprimé en  $U\sqrt{s}$  ou en  $U/\sqrt{\text{Hz}}$ , c'est-à-dire en  $\text{deg/s}/\sqrt{s}$  ou en  $\text{deg/s}/\sqrt{\text{Hz}}$  (parfois en  $\text{deg}/\sqrt{\text{heure}}$ ).

Dans une simulation numérique utilisant un pas de calcul  $T = \Delta t$ , si on veut simuler ce bruit à l'aide d'un générateur aléatoire numérique fournissant des valeurs aléatoires  $\eta$  à moyenne nulle et à écart-type unité, les valeurs numériques successives de  $b'_\omega(k\Delta t)$  seront obtenues en multipliant chaque tirage d'une valeur  $\eta$  par  $\frac{\sigma_\omega}{\sqrt{\Delta t}}$ .

Ajoutons à notre exemple une dérive aléatoire  $d(t)$  obtenue par l'intégration d'un bruit blanc. Le modèle devient :

$$\begin{aligned}\omega_m(t) &= \omega(t) + d(t) + b_\omega(t) \\ d'(t) &= b_a(t)\end{aligned}$$

où  $b_a(t)$  est, par exemple, un bruit blanc d'origine électronique. Cette deuxième équation appelée *marche au hasard*, s'apparente au cas traité en bruit d'état. Comme  $d(t)$  s'exprime en  $U = \text{deg/s}$ , le bruit  $b_a(t)$  va être caractérisé par son pseudo écart-type  $\sigma_{b_a} = \sqrt{Q_c}$  exprimé en  $U/\sqrt{s}$  ou en  $(U/s)/\sqrt{\text{Hz}}$ , c'est-à-dire en  $\text{deg/s}/\sqrt{s}$  ou en  $(\text{deg/s}^2)/\sqrt{\text{Hz}}$  (parfois les heures remplacent les secondes).

Si on simule cette marche au hasard en numérique en tirant une nouvelle valeur aléatoire  $\xi$  à moyenne nulle et à écart-type unité à chaque pas de calcul  $\Delta t$ , les valeurs successives de  $b'_a(k\Delta t)$  seront obtenues en multipliant chaque tirage la valeur  $\xi$  par  $\frac{\sigma_{b_a}}{\sqrt{\Delta t}}$ . Si l'intégration est effectuée par une simple méthode du 1er ordre du type  $d((k+1)\Delta t) = d(k\Delta t) + b'_a(k\Delta t)\Delta t$ , on aura alors  $d((k+1)\Delta t) = d(k\Delta t) + \xi \frac{\sigma_{b_a}}{\sqrt{\Delta t}}\Delta t = d(k\Delta t) + \xi \sigma_{b_a} \sqrt{\Delta t}$ , où on voit que la quantité ajoutée à  $d(k\Delta t)$  a bien la même unité (à savoir  $U = \text{deg/s}$ ).



## 6.4 Régimes permanents des filtres

L'examen des limites atteintes en régime permanent permet de qualifier les propriétés du filtre de Kalman.

### 6.4.1 Filtre continu

On a :

$$\dot{\Lambda}_c = Q_c - \frac{\Lambda_c^2}{R_c}$$

qui a pour régime permanent ( $\dot{\Lambda}_c = 0$ ) :

$$\Lambda_c = \sqrt{R_c Q_c}$$

La covariance d'erreur d'estimation tend vers la moyenne géométrique des covariances de bruit de mesure et de bruit d'état en continu.

### 6.4.2 Filtre numérique

On a :

$$\Lambda_{dn+1}^+ = \frac{R_d (\Lambda_{dn}^+ + Q_d)}{R_d + \Lambda_{dn}^+ + Q_d}$$

qui a pour régime permanent la solution de :

$$\Lambda_d^2 + Q_d \Lambda_d - Q_d R_d = 0$$

soit :

$$\Lambda_d = \sqrt{Q_d R_d + \frac{Q_d^2}{4}} - \frac{Q_d}{2}$$

Si on veut que  $\Lambda_d \simeq \Lambda_c$ , c'est-à-dire que les qualités des deux filtres soient voisines, on doit avoir :

$$\sqrt{R_c Q_c} \simeq \sqrt{Q_d R_d + \frac{Q_d^2}{4}} - \frac{Q_d}{2}$$

Or  $Q_d \simeq Q_c T$ , d'où :

$$\sqrt{R_c Q_c} \simeq \sqrt{Q_c R_d T + \frac{T^2 Q_c^2}{4}} - \frac{T Q_c}{2}$$

Cette relation sera d'autant plus vraie que  $R_d T \simeq R_c$  (acquis a priori) et que  $T$  est petit.

En conclusion, les relations :

$$\begin{aligned} Q_d &\simeq Q_c T \\ R_d T &\simeq R \\ &\text{et } T \text{ petit} \end{aligned}$$

sont à rechercher si on désire que les filtres continu et numérique aient le même comportement.

## 6.5 Annexe

Montrons par un calcul approximatif que lorsque :

$$\dot{x}(t) = v(t)$$

avec  $v(t)$  bruit blanc tel que :

$$E[v(t)v(t+\tau)] = R(\tau) = a\delta(\tau)$$

la dérivée de la covariance de  $x(t)$  vaut  $a$ .

Calculons cette dérivée :

$$\Lambda_c(t) = E[x^2(t)]$$

Il vient (des approximations, rien que des approximations !!!):

$$\begin{aligned} \dot{\Lambda}_c(t) &= 2E[\dot{x}(t)x(t)] \\ &= 2E[v(t)x(t)] \\ &= 2E\left[v(t)\int_0^t \dot{x}(u)du\right] \\ &= 2E\left[v(t)\int_0^t v(u)du\right] \\ &= 2\int_0^t E[v(t)v(u)]du \\ &= 2a\int_0^t \delta(u-t)du \\ &= 2a\int_{-t}^0 \delta(\tau)d\tau \end{aligned}$$

Si on considère que  $\int_{-t}^t \delta(\tau)d\tau = 1 = \int_{-t}^0 \delta(\tau)d\tau + \int_0^t \delta(\tau)d\tau = 2\int_{-t}^0 \delta(\tau)d\tau$  on a finalement :

$$\dot{\Lambda}_c(t) = a$$