

# **THÉORIE DES ERREURS :**

**ESTIMATION DE  $r$  GRANDEURS INCONNUES  
À PARTIR DE  $n$  GRANDEURS OBSERVÉES  
PAR LA MÉTHODE DES MOINDRES CARRÉS**

**PAR**

**PHILIPPE HOTTIER, INGÉNIEUR EN CHEF  
GÉOGRAPHE**

**- IGN - 1980 -**

**ÉDITION NUMÉRIQUE RÉALISÉE PAR**

**ABDELMAJID BEN HADJ SALEM  
INGÉNIEUR GÉNÉRAL GÉOGRAPHE**

**Août 2020**

**version 1.**

**Abdelmajid BEN HADJ SALEM**

e-mail : abenhadsalem@gmail.com

*À mon professeur Philippe  
Hottier*

*À ma femme, à mes enfants*

*À la mémoire de mon ami et  
collègue Mohamed Rached Boussema<sup>1</sup>*

---

1. Décédé le 20 juillet 2020.



# Table des matières

|          |  |    |
|----------|--|----|
| <b>1</b> | <b>Introduction - Poids</b> .....  | 1  |
| 1.1      | DÉFINITION DU PROBLÈME .....   | 1  |
| 1.1.1    | Définition du problème .....   | 1  |
| 1.1.2    | Les Mesures .....  | 2  |
| 1.1.3    | Système S de mesures .....   | 3  |
| 1.2      | EXEMPLE FONDAMENTAL .....  | 5  |
| 1.3      | POIDS - MATRICE DE POIDS - VARIANCE DE LA MESURE DE<br>POIDS UNITAIRE .....  | 6  |
| 1.4      | REMARQUE : NOMBRE DE MESURES PAR GRANDEUR OBSERVÉE   | 8  |
| <b>2</b> | <b>Calcul des Estimations des <math>r</math> Grandeurs Inconnues par les<br/>Moindres Carrés</b> .....   | 11 |
| 2.1      | NOTATIONS .....  | 12 |
| 2.2      | POSE DES EQUATIONS .....   | 13 |
| 2.2.1    | Méthode des équations d'observations .....   | 14 |
| 2.2.2    | Méthode des équations de condition liant les<br>$n$ grandeurs observées ( $r$ grandeurs inconnues<br>indépendantes; $r$ = nombre de degrés de liberté de la<br>figure ou du phénomène) ..... | 22 |

|          |  |           |
|----------|--|-----------|
| 2.3      | DÉTERMINATION D'UN ESTIMATEUR $\tilde{X}$ DE $\dot{X}$ PAR LA MÉTHODE DES MOINDRES CARRÉS - ESTIMATEUR D'UNE FONCTION LINÉAIRE DE $\dot{X}$ - ESTIMATEURS DE $\dot{L}$ ET $\dot{V}$ - INVARIANCE DE LA SOLUTION PAR CHANGEMENT D'INCONNUES . . . . . | 28        |
| 2.3.1    | Problème linéaire . . . . .  | 28        |
| 2.3.2    | Détermination de l'estimateur $\tilde{X}$ de $\dot{X}$ . . . . .   | 29        |
| 2.3.3    | Estimateur d'une fonction linéaire de $\dot{X}$ . . . . .  | 35        |
| 2.3.4    | Estimateurs de $\dot{L}$ et $\dot{V}$ . . . . .  | 36        |
| 2.3.5    | Invariance de la solution par changement d'inconnues . . . . .   | 37        |
| 2.3.6    | Equivalence d'un système d'équations d'observations (avec éventuellement équations de condition liant les inconnues) et d'un système d'équations de condition liant les grandeurs observées . . . . .  | 39        |
| 2.4      | ESTIMATEUR AU SENS DES MOINDRES CARRÉS D'UNE FONCTION DES INCONNUES $\dot{Y} = f(\dot{X})$ . . . . .   | 40        |
| 2.5      | PRATIQUE DES CALCULS . . . . .   | 41        |
| 2.5.1    | Retour sur la méthode des équations de condition liant les grandeurs observées . . . . .   | 41        |
| 2.5.2    | Compensation par les éléments . . . . .  | 42        |
| 2.5.3    | Formulaire . . . . .   | 43        |
| 2.5.4    | Rappel sur les méthodes d'inversion d'un système linéaire régulier . . . . .   | 46        |
| <b>3</b> | <b>Compléments Sur La Théorie Élémentaire Des Moindres Carrés .</b>  | <b>63</b> |
| 3.1      | FORME CANONIQUE D'UN SYSTÈME D'ÉQUATIONS D'OBSERVATIONS . . . . .  | 64        |
| 3.2      | ESTIMATION $\tilde{X}$ D'UNE GRANDEUR À $r$ DIMENSIONS $\dot{X}$ À PARTIR D'UN ECHANTILLON DE $N$ ESTIMATIONS $X_1, X_2, \dots, X_N$ DE POIDS $P_1, P_2, \dots, P_N$ ( $P_i = \text{matrice}$ ) . . . . .  | 65        |
| 3.3      | CONTRIBUTIONS DE CHAQUE EQUATION D'OBSERVATIONS, OU D'UNE PARTIE DU SYSTÈME DES EQUATIONS D'OBSERVATIONS, À LA MATRICE NORMALE DANS LE CAS OÙ LES OBSERVATIONS SONT INDÉPENDANTES . . . . .  | 66        |

|          |  |            |
|----------|--|------------|
| 3.4      | PROBLÈME DES OBSERVATIONS RÉPÉTÉES- CHOIX DES POIDS POUR UN SYSTÈME DE MESURES ENTIÈREMENT DÉTERMINÉ- CHOIX DES POIDS POUR DES OBSERVATIONS SUPPLÉMENTAIRES D'UNE GRANDEUR DÉJÀ OBSERVÉE . . . . . | 67         |
| 3.4.1    | Groupement des observations relatives à une même grandeur observée . . . . .   | 67         |
| 3.4.2    | Etude d'un système de mesures entièrement déterminé  | 70         |
| 3.4.3    | Choix des poids d'observations supplémentaires relatives à une grandeur observée . . . . .   | 71         |
| 3.5      | ADJONCTION D'OBSERVATIONS OU D'INCONNUES SUPPLÉMENTAIRES (cas de grands réseaux) - MÉTHODE DU BORDAGE . . . . .  | 73         |
| <b>4</b> | <b>Propriétés Statistiques des Estimateurs des Moindres Carrés . . . .</b>   | <b>79</b>  |
| 4.1      | INTRODUCTION . . . . .   | 80         |
| 4.2      | PROPRIÉTÉS STATISTIQUES DES ESTIMATEURS DES MOINDRES CARRÉS DANS L'HYPOTHÈSE OÙ LES ERREURS DE MESURES SONT ACCIDENTELLES- VARIANCE DES ESTIMATEURS . . . . .                                      | 81         |
| 4.2.1    | Les estimateurs des moindres carrés . . . . .  | 81         |
| 4.2.2    | Estimateur de la variance unitaire $\sigma_0^2$ . . . . .  | 82         |
| 4.2.3    | Expression des matrices-variance des inconnues, des observations compensées, des résidus, de toute fonction linéaire des inconnues . . . . .   | 85         |
| 4.3      | PROPRIÉTÉS STATISTIQUES DES ESTIMATEURS DES MOINDRES CARRÉS DANS L'HYPOTHÈSE OÙ LES ERREURS DE MESURES SONT NORMALES ET CENTRÉES . . . . .   | 90         |
| 4.4      | COMPARAISON DE LA VARIANCE UNITAIRE A PRIORI $\sigma_0^2$ QUAND ELLE EST CONNUE AVEC LA VARIANCE UNITAIRE ESTIMÉE $s_0^2$ . .  | 100        |
| <b>5</b> | <b>Précision, Justesse et Exactitude d'un Ensemble de Grandeurs Estimées par la Méthode des Moindres Carrés . . . . .</b>  | <b>103</b> |
| 5.1      | INTRODUCTION . . . . .   | 104        |
| 5.1.1    | Définition de l'erreur . . . . .   | 105        |

|       |  |            |
|-------|--|------------|
| 5.2   | ESTIMATIONS PONCTUELLES DE LA PRÉCISION, DE LA JUSTESSE ET DE L'EXACTITUDE DE L'ESTIMATEUR D'UNE GRANDEUR A $r$ DIMENSIONS DANS LE CADRE D'UN SYSTÈME $S$ DE MESURES . . . . .   | 106        |
| 5.2.1 | Précision de l'estimateur $\tilde{X}$ . . . . .  | 106        |
| 5.2.2 | Justesse de l'estimateur $\tilde{X}$ . . . . .   | 108        |
| 5.2.3 | Exactitude de l'estimateur $\tilde{X}$ . . . . .   | 109        |
| 5.3   | ESTIMATION PONCTUELLE DE LA PRÉCISION DE LA VARIANCE ESTIMÉE ET DE L'ECART-TYPE ESTIMÉ DE LA MESURE DE POIDS UNITAIRE (DANS LE CAS OÙ LA REDONDANCE $n - r$ EST FORTE) . . . . . | 111        |
| 5.4   | ESTIMATION DE LA PRÉCISION ET DE L'EXACTITUDE DES ESTIMATEURS PAR LA MÉTHODE DES INTERVALLES DE CONFIANCE - INTERVALLE DE CONFIANCE POUR L'ECART-TYPE UNITAIRE . . . . .         | 112        |
| 5.4.1 | Intervalle de confiance pour l'espérance de l'estimateur d'une grandeur inconnue scalaire dans le cadre d'un certain système $S$ de mesures . . . . .                            | 113        |
| 5.4.2 | Intervalle de confiance pour la valeur universelle d'une grandeur inconnue scalaire . . . . .  | 113        |
| 5.4.3 | Intervalle de confiance pour l'écart-type unitaire $\sigma_0$ . . . . .  | 114        |
| 5.5   | LES APPLICATIONS LINÉAIRES . . . . .   | 115        |
| 5.6   | OPÉRATIONS SUR LES MATRICES . . . . .  | 116        |
| 5.7   | PROPRIÉTÉS DES MATRICES . . . . .  | 117        |
| 5.8   | VALEURS PROPRES D'UNE MATRICE . . . . .  | 119        |
| 5.8.1 | Vecteurs propres . . . . .   | 119        |
| 5.9   | RANG D'UNE APPLICATION LINÉAIRE . . . . .  | 120        |
| 5.10  | EXERCICES . . . . .  | 120        |
|       | <b>LISTE DES FIGURES . . . . .</b>   | <b>126</b> |
|       | Liste des Figures . . . . .  | 126        |



# Préface

C'est avec un grand plaisir que je vous présente cette première version numérique de la polycopie du cours " Théorie des erreurs " de l'Ingénieur en Chef Géographe Philippe Hottier.

Ce cours a été enseigné au début des années 80 du dernier siècle à l'Ecole Nationale des Sciences Géographiques (ENSG) de l'Institut Géographique National de France, pour les élèves ingénieurs des cycles A et B.

J'étais son élève en première année du cycle A de l'année universitaire 1984-1985 où j'avais suivi son cours sur la théorie des erreurs.

Cette version numérique a gardé quasiment le texte original avec :

- ajout d'un sommaire pour chaque chapitre,
- présentation des conclusions, les remarques et les résultats importants avec un arrière plan grisâtre,
- présentation d'une démonstration du théorème de Linnik (pages 95-97),
- ajout d'une annexe comme rappels sur le calcul matriciel,
- ajout d'une deuxième annexe comportant une table de la loi du  $\chi^2$ .

Dans l'ensemble, cette version numérique est un apport aux étudiants en géomatique dans le domaine de la théorie des erreurs. Tout commentaire est la bienvenue.

Nous saisissons cette opportunité de présenter à travers cette publication nos remerciements et sincères salutations à Monsieur Philippe Hottier.

Lors de la finition de ce document, la Tunisie a perdu un des grands spécialistes internationaux en télédétection et systèmes d'information géographique mon ami et collègue le Docteur-Ingénieur Mohamed Rached Boussema, ancien Professeur de Géomatique à l'Ecole Nationale des Ingénieurs de Tunis. Je dédie cette édition numérique à sa mémoire.

Tunis,  
Août 2020

*A. Ben Hadj Salem*  
*Ingénieur Général Géographe*



**Fig. 0.1** Prof. Mohamed Rached Boussema (1953-2020)

# Chapitre 1

## Introduction - Poids

### Sommaire

---

|            |   |          |
|------------|---|----------|
| <b>1.1</b> | <b>DÉFINITION DU PROBLÈME</b> .....   | <b>1</b> |
| 1.1.1      | Définition du problème .....  | 1        |
| 1.1.2      | Les Mesures.....  | 2        |
| 1.1.3      | Système S de mesures.....   | 3        |
| <b>1.2</b> | <b>EXEMPLE FONDAMENTAL</b> .....  | <b>4</b> |
| <b>1.3</b> | <b>POIDS - MATRICE DE POIDS - VARIANCE DE LA MESURE DE<br/>POIDS UNITAIRE</b> ..... | <b>6</b> |
| <b>1.4</b> | <b>REMARQUE : NOMBRE DE MESURES PAR GRANDEUR OBSERVÉE</b>                           | <b>8</b> |

---

### 1.1 DÉFINITION DU PROBLÈME

#### 1.1.1 Définition du problème

On veut déterminer  $r$  grandeurs scalaires inconnues  $G_1, G_2, \dots, G_r$  à l'aide de  $n$  grandeurs observées distinctes ou non des précédentes, mais qui leur sont liées géométriquement (ou physiquement).

On suppose qu'il n'a qu'une seule mesure par grandeur, étant entendu qu'il peut s'agir d'une mesure résumée (voir remarque ci-dessous § 4). Il y a donc  $n$  mesures (ou observations)  $L_1, L_2, \dots, L_n$  : généralement le nombre des mesures est surabondant par rapport au nombre des inconnues à déterminer,

et il conviendra de "compenser" les mesures et d'estimer au mieux les inconnues.

Les grandeurs à déterminer peuvent être ou non liées entre elles; s'il y'a  $p$  relations entre ces grandeurs, nous dirons que le nombre de degrés de liberté de l'ensemble est  $r - p$  : et il n'y aura en fait que  $r - p$  inconnues scalaires. Une condition nécessaire pour qu'on puisse déterminer ces  $r - p$  inconnues est que  $n \geq r - p$ , mais elle n'est pas évidemment pas suffisante : il faut en outre que parmi les  $n$  mesures, il s'en trouve au moins un groupe permettant de calculer les inconnues : autrement dit il faut que les grandeurs observées déterminent géométriquement (ou physiquement) les grandeurs à déterminer.

### 1.1.2 Les Mesures

Nous supposons que chacune des  $n$  grandeurs observées l'a été selon un processus dont l'exactitude est bien déterminée (mais non nécessairement connue) : autrement dit, lorsque nous voudrions répéter la détermination des  $r$  inconnues à partir d'une autre série de  $n$  mesures des  $n$  grandeurs observées, nous supposons ces  $n$  nouvelles mesures faites dans des conditions d'exactitude inchangées (mais le processus, c'est-à-dire les appareils, les observateurs..., peuvent eux changer).

Les mesures  $l_1, l_2, \dots, l_n$  condensées dans le vecteur :

$${}_n L_1 = \begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix} \quad (1.1)$$

peuvent être :

- directes, et dans ce cas indépendantes; outre leur valeur, nous n'en connaissons que l'exactitude, chiffrée par leur erreur moyenne quadratique, et nous supposons qu'elles sont les valeurs de variables aléatoires normales :

$$L_i \in \mathcal{N}(\dot{L}_i, emq_i)$$

$\dot{L}_i$  étant la valeur universelle de la grandeur observée correspondante (définie comme moyenne d'un grand nombre de mesures de cette grandeur, effectuées avec un grand nombre de processus de mesure de caractéristiques analogues).

La matrice variance du vecteur observation  ${}_nL_1$  est alors diagonale.

$$\Gamma_L = \begin{pmatrix} emq_1^2 & 0 & 0 & 0 \\ 0 & emq_2^2 & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & emq_n^2 \end{pmatrix} \quad \text{et} \quad {}_nL_1 \in \mathcal{N}(\dot{L}, \sqrt{\Gamma_L}) \quad (1.2)$$

- indirectes, c'est-à-dire résultant de calculs faits avec des mesures directes en amont de ces calculs, elles sont alors en principe généralement corrélées, corrélation qui se traduit par le fait que  $\Gamma_L$  n'est pas diagonale.

Le vecteur  ${}_nL_1$  sera encore considéré comme normal :

$${}_nL_1 \in \mathcal{N}(\dot{L}, \sqrt{\Gamma_L})$$

Autrement dit la densité de probabilité d'un tel vecteur sera :

$$\frac{1}{\sqrt{(2\pi)^n |\text{Dét}(\Gamma_L)|}} e^{-\frac{1}{2} (L-\dot{L})^T \Gamma_L^{-1} (L-\dot{L})}$$

Dans tout ce qui suit, nous supposons donc en principe les mesures normales, nous savons en effet que si on emploie des instruments bien conçus, cette hypothèse est toujours acceptable; elle l'est d'autant plus que les observations qu'on envisage ici sont très souvent en pratique des moyennes d'observations élémentaires (et la loi des grands nombres nous dit que la moyenne d'un nombre suffisant d'observations est sensiblement normale).

Dans le cas où les  $n$  mesures ne sont pas normales, nous verrons d'ailleurs que l'essentiel de la théorie des moindres carrés subsiste.

### 1.1.3 Système S de mesures

Il comprend par définition :

- les  $r$  grandeurs inconnues,

- les  $n$  grandeurs observées dans des conditions d'exactitude déterminée.

Répéter une détermination dans le cadre du système  $S$ , consistera alors à mesurer les mêmes  $r$  grandeurs dans les mêmes conditions d'exactitude, et à évaluer les valeurs des grandeurs inconnues par la même méthode.

Le but que nous nous proposons est alors :

- d'estimer au mieux à partir de la donnée des  $n$  mesures la valeur des  $r$  grandeurs inconnues, ou encore étant donné un échantillon aléatoire  $(L_1, L_2, \dots, L_n)$  de déterminer le meilleur estimateur des valeurs des grandeurs inconnues.

Nous appellerons :

$${}_r\tilde{\mathcal{X}}_1 = \begin{pmatrix} \tilde{\mathcal{X}}_1 \\ \tilde{\mathcal{X}}_2 \\ \vdots \\ \tilde{\mathcal{X}}_r \end{pmatrix} \quad (1.3)$$

le vecteur des estimateurs correspondants.  $\tilde{\mathcal{X}}$  sera, en un sens qui sera précisé, le meilleur estimateur de  $E(\tilde{\mathcal{X}})$  espérance de toutes les estimations qu'on peut obtenir dans le cadre du système  $S$  de mesures (on aura  $E(\tilde{\mathcal{X}}) \neq \tilde{\mathcal{X}}$ , nous savons qu'il n'est pas possible d'atteindre  $\tilde{\mathcal{X}}$  dans le cadre du système  $S$ , qui n'est que l'un des systèmes de mesures qu'il est possible de mettre en œuvre pour déterminer les grandeurs inconnues).

- de trouver le meilleur estimateur de la matrice variance du vecteur observation  $L$ , c'est-à-dire  $\Gamma_L$ ; cela n'est évidemment à priori possible que sous certaines hypothèses : par exemple dans le cas où les mesures sont directes et indépendantes, si on peut les regrouper en catégories de mesures de même exactitude, chaque catégorie comporte au moins 2 mesures.

- de trouver le meilleur estimateur de la matrice variance  $\Gamma_{\tilde{\mathcal{X}}}$  de  $\tilde{\mathcal{X}}$  qui chiffrera "la précision" de  $\tilde{\mathcal{X}}$  et d'évaluer l'exactitude de  $\tilde{\mathcal{X}}$  si se peut faire.

## 1.2 EXEMPLE FONDAMENTAL

Dans toute la suite, nous nous référons pour fixer les idées, et pour les applications numériques à l'exemple suivant, qui sera dit exemple fondamental.

Soit un triangle de côtés  $a, b, c$  et d'angles  $A, B$  et  $C$ .

On se propose :

- d'estimer  $a, b$  et  $c$ , ainsi que d'estimer la précision et l'exactitude de ces déterminations,
- d'estimer l'aire du triangle, ainsi que la précision et l'exactitude de l'estimation, à partir des données suivantes :

$$\left\{ \begin{array}{l} a = 96.48 \text{ mm} \quad (\text{moyenne de 6 mesures élémentaires} \\ \quad \quad \quad \text{au double-écimètre}), \\ b = 115.50 \text{ mm} \quad (\text{une seule mesure élémentaire}), \\ A = 63.042 \text{ gr} \\ B = 99.802 \text{ gr} \\ C = 37.008 \text{ gr} \end{array} \right. \quad (1.4)$$

Les angles sont obtenues par la moyenne de 6 mesures élémentaires, avec :

- l'écart-type d'une mesure élémentaire de distance :  $\sigma_d = 0.18 \text{ mm}$ ,
- l'écart-type d'une mesure élémentaire d'angle :  $\sigma_a = 0.075 \text{ gr}$ .

Il est clair que le problème est bien déterminé géométriquement et que les données sont surabondantes (5 mesures pour 3 inconnues à 3 degrés de liberté).

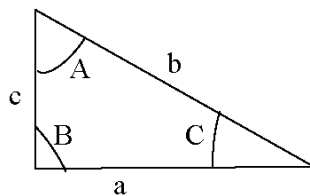


Fig. 1.1 L'Exemple Fondamental

### 1.3 POIDS - MATRICE DE POIDS - VARIANCE DE LA MESURE DE POIDS UNITAIRE

Le problème qu'on s'est posé, à savoir déterminer à la fois la meilleure estimation de  $E(\mathcal{X})$ , et la variance de plusieurs catégories d'observations est inextricable sauf :

- si on suppose que  $\Gamma_L$ , matrice variance chiffrant l'exactitude de  $L$  est connue; cas très fréquent; on connaît généralement, par exemple dans le cas de mesures directes et indépendantes, l'exactitude des diverses sortes d'observations,
- si on suppose, cas un peu plus général que le précédent, que  $\Gamma_L$  matrice variance chiffrant l'exactitude des observations est connue à un facteur multiplicatif près, que la théorie permettra alors d'estimer.

Nous nous placerons toujours dans le second cas, ce qui nous permettra si nous sommes dans le premier, de disposer d'un test permettant de vérifier la validité des données sur l'exactitude de diverses mesures.

Nous poserons donc :

$$\Gamma_L = \frac{\sigma_0^2}{P} = \sigma_0^2 P^{-1} \text{ avec } \begin{cases} P = \text{matrice de poids (fixée avant les calculs)} \\ \sigma_0^2 = \text{scalaire inconnu à estimer la variance unitaire} \end{cases} \quad (1.5)$$

La matrice des poids doit donc être choisie comme inversement proportionnelle à la matrice variance des erreurs de mesures.

En particulier, si les mesures sont indépendantes, les poids ( $P$  est alors diagonale et les poids sont les termes diagonaux) doivent être choisis comme inversement proportionnels aux carrés des erreurs moyennes quadratiques ( $emq$ ).

Une fois  $P$  fixée, on a défini alors "de facto" la variance de la mesure de poids unitaire  $\sigma_0^2$  que nous désignerons en abrégé par variance unitaire.

Dans le cas où on connaît l'exactitude du vecteur observation  $L$ , on connaît à priori la valeur de cette variance, valeur que nous appellerons alors variance unitaire



à priori.

Très souvent d'ailleurs, si  $\Gamma_L$  est connu, on choisit  $P = \Gamma_L^{-1}$ , ce qui conduit à considérer une variance unitaire à priori égale à 1 ; dans le cas de mesures directes et indépendantes, les poids sont alors les inverses des carrés des  $emq$ .

Une comparaison de la variance unitaire à priori  $\sigma_0^2$ , avec la valeur estimée  $s_0^2$  permettra de déceler des incohérences dans les données relatives à l'exactitude des mesures (entre autres).

#### **Application à l'Exemple Fondamental - Unités normalisées**

Le vecteur d'observation  $L$  aura pour composantes  $a, b, A, B$  et  $C$ . Nous prendrons des poids égaux aux inverses des carrés des  $emq$  de chaque observation.

Mais nous ferons d'abord un changement d'unité pour chaque type d'observation, les nouvelles unités que nous appellerons unités normalisées étant choisies de telle sorte que les  $emq$  de chaque type d'observation s'expriment par des nombres voisins de l'unité.

Cette précaution est essentielle dans les calculs manuels, ou semi-manuels quand il faut noter (avec un nombre limité de chiffres significatifs) des résultats intermédiaires ; on verra en effet qu'elle entraîne la normalisation de tous les coefficients qui interviennent dans les calculs ultérieurs.

Nous choisirons ici comme unité normalisées :

- le décimillimètre pour les mesures de distances,
- le décimilligrade pour les angles.

Dans ces conditions, on adoptera pour  $emq$  des diverses mesures les chiffres suivants :

$$emq_a = \sqrt{1 + 1/6} \times 1.8 \approx 1.9 \Rightarrow p_a = 1/1.9^2 = 0.277$$

$$emq_b = \sqrt{2} \times 1.8 \approx 2.5 \Rightarrow p_b = 1/2.5^2 = 0.160$$

$$emq_{A,B,C} = \sqrt{1 + 1/6} \times 0.8 \approx 0.81 \Rightarrow p_{A,B,C} = 1/0.81^2 = 1.524$$

(On a supposé pour évaluer l'exactitude que le biais était de l'ordre de l'écart-type d'une mesure élémentaire).

On aura donc, puisque les mesures sont indépendantes :

$$P = \begin{pmatrix} 0.277 & 0 & 0 & 0 & 0 \\ 0 & 0.160 & 0 & 0 & 0 \\ 0 & 0 & 1.524 & 0 & 0 \\ 0 & 0 & 0 & 1.524 & 0 \\ 0 & 0 & 0 & 0 & 1.524 \end{pmatrix} \quad \text{avec} \quad \begin{cases} \text{variance unitaire à priori } \sigma_0^2 = 1 \\ \text{unité de distance : } 0.1 \text{ mm} \\ \text{unité d'angle : } 10 \text{ minutes centé} \end{cases} \quad (1.6)$$

On notera qu'une fois fixées les  $emq$ , la précision avec laquelle doivent être calculés les poids n'a pas besoin d'être très grande : il convient simplement que toutes les quantités  $p_i emq_i^2$  ( $i = 1, n$ ) soient alors égales par exemple à 1% près, de façon à ne pas fausser au départ la valeur de la variance unitaire à priori.

Ainsi, dans le cas présent, si pour les angles on prend au lieu de 1.524 le poids 1.520, on aura :

$$1.520 \times 0.81^2 = 0.997$$

ce poids convient tout aussi bien que le précédent.

#### 1.4 REMARQUE : NOMBRE DE MESURES PAR GRANDEUR OBSERVÉE

Nous avons supposé qu'il n'y avait qu'une mesure par grandeur observée, étant entendu qu'il peut s'agir d'une mesure "résumée" (moyenne pondérée ou non de plusieurs mesures) dont le poids est calculé en fonction de l'exactitude et non de la précision de cette mesure résumée.

Les raisons en sont les suivantes :

\* pour le choix de l'exactitude, et non de la précision dans le calcul du poids de l'observation résumée. C'est toujours ce qui est fait dans la pratique.

Pour éclairer la chose considérons la situation caricaturale suivante :

On détermine un point  $M$  par intersection depuis deux stations  $S_1$  et  $S_2$ , en effectuant (avec des instruments de même exactitude) 10000 visées depuis  $S_1$ , et une seule depuis  $S_2$ .

Si on suivait alors certains exposés classiques, il faudrait alors introduire dans les calculs la visée résumée issue de  $S_1$  avec le poids 100, et celle issue de  $S_2$  avec le poids 1 en se basant sur le fait que la variance de la première est 100 fois plus faible que celle de la seconde.

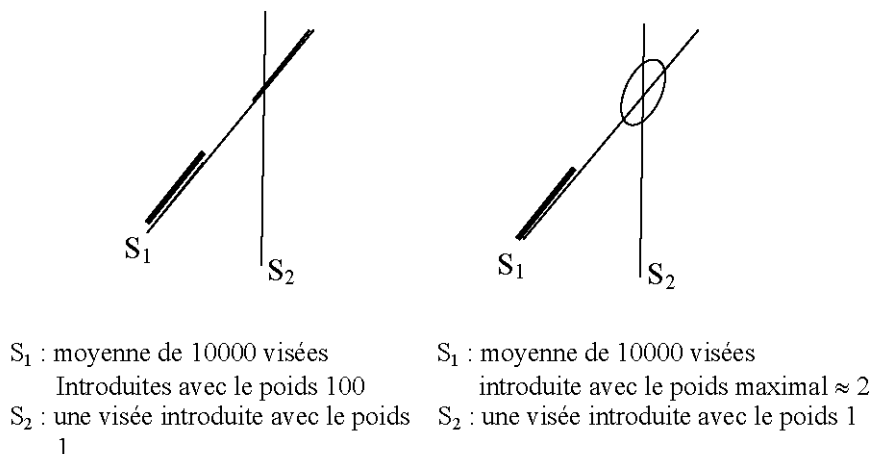
En pratique, sauf cas tout à fait exceptionnel (par exemple étude de la fidélité d'un système spécifié jusque dans ses moindres détails) on ne procédera jamais ainsi.

On accordera certes un poids un peu plus fort (à notre avis le poids 2, pour le poids 1 à la visée issue de  $S_2$ ) à la visée résumée issue de  $S_1$ , mais jamais dans le rapport 100 à 1.

La raison en est que ne compte en pratique que l'exactitude obtenue pour l'estimation de  $M$  et que cette exactitude ne dépend que de l'angle sous lequel se coupent les 2 visées, et des exactitudes des 2 visées, non de leur précision. Or, l'ellipse indicatrice de la matrice variance obtenue à partir de poids évalués en fonction des précisions, se réduit ici pratiquement à un segment porté par la visée résumée de  $S_1$ , et ne permet pas d'évaluer l'exactitude de la détermination.

Mais dans le cas où on prend des poids évalués en fonction des exactitudes, on obtient une matrice  $\Gamma$  non singulière et qui permet elle, d'évaluer raisonnablement l'exactitude de la détermination de  $M$  (on pourra si on admet que le biais est chiffré par la dispersion des déterminations élémentaires équivalentes, évaluer la variance des erreurs sur  $M$  par  $\Gamma + \frac{n}{r}\Gamma \approx 2\Gamma$ ).

\* pour la nécessité de ne prendre qu'une seule mesure résumée par grandeur observée : il n'y a pas de relation simple entre l'exactitude d'une mesure résumée, et l'exactitude des mesures constituantes, donc entre le poids de celles-ci et le poids de celle-là.



**Fig. 1.2** Le Choix des poids

Or, on voit facilement que pour que la théorie des moindres carrés donne mêmes résultats que l'on introduise plusieurs mesures par grandeur observée, ou que l'on en introduise qu'une, résumée des mesures précédentes, il faut que le poids de la mesure résumée soit égal à la somme des poids des mesures constituantes; ce qui n'est vrai si on calcule les poids en fonction des exactitudes.

C'est pourquoi nous conviendrons dans la suite qu'il n'y a qu'une seule mesure par grandeur observée : cela n'offre guère d'inconvénient en pratique.

Il existe cependant certains problèmes pour lesquels il est commode de pouvoir envisager plusieurs mesures pour une seule grandeur observée (par exemple adjonction après coup d'observations supplémentaires); dans ce cas nous serons conduits à ajouter de telles observations supplémentaires avec des poids dégressifs de telle sorte que l'exactitude de l'observation résumée correspondante, soit celle de l'observation résumée qu'on aurait autrement introduite.

# Chapitre 2

## Calcul des Estimations des $r$ Grandeurs Inconnues par les Moindres Carrés

### Sommaire

---

|       |  |    |
|-------|--|----|
| 2.1   | NOTATIONS .....  | 12 |
| 2.2   | POSE DES EQUATIONS .....   | 13 |
| 2.2.1 | Méthode des équations d'observations .....   | 13 |
| 2.2.2 | Méthode des équations de condition liant les $n$ grandeurs observées ( $r$ grandeurs inconnues indépendantes; $r$ = nombre de degrés de liberté de la figure ou du phénomène) .....  | 22 |
| 2.3   | DÉTERMINATION D'UN ESTIMATEUR $\hat{X}$ DE $X$ PAR LA MÉTHODE DES MOINDRES CARRÉS - ESTIMATEUR D'UNE FONCTION LINÉAIRE DE $X$ - ESTIMATEURS DE $\hat{L}$ ET $\hat{V}$ - INVARIANCE DE LA SOLUTION PAR CHANGEMENT D'INCONNUES | 27 |
| 2.3.1 | Problème linéaire .....  | 27 |
| 2.3.2 | Détermination de l'estimateur $\hat{X}$ de $X$ .....   | 29 |
| 2.3.3 | Estimateur d'une fonction linéaire de $X$ .....  | 34 |
| 2.3.4 | Estimateurs de $\hat{L}$ et $\hat{V}$ .....  | 36 |
| 2.3.5 | Inviance de la solution par changement d'inconnues .....   | 36 |

|            |   |           |
|------------|---|-----------|
| 2.3.6      | Equivalence d'un système d'équations d'observations (avec éventuellement équations de condition liant les inconnues) et d'un système d'équations de condition liant les grandeurs observées ..... | 38        |
| <b>2.4</b> | <b>ESTIMATEUR AU SENS DES MOINDRES CARRÉS D'UNE FONCTION DES INCONNUES <math>\mathcal{Y} = f(\mathcal{X})</math></b> .....  | <b>39</b> |
| <b>2.5</b> | <b>PRATIQUE DES CALCULS</b> .....   | <b>40</b> |
| 2.5.1      | Retour sur la méthode des équations de condition liant les grandeurs observées .....  | 40        |
| 2.5.2      | Compensation par les éléments .....   | 42        |
| 2.5.3      | Formulaire .....  | 43        |
| 2.5.4      | Rappel sur les méthodes d'inversion d'un système linéaire régulier .....  | 46        |

## 2.1 NOTATIONS

${}_r\dot{\mathcal{X}}_1 = \begin{pmatrix} \dot{\mathcal{X}}_1 \\ \dot{\mathcal{X}}_2 \\ \vdots \\ \dot{\mathcal{X}}_r \end{pmatrix}$

{ vecteur dont les composantes sont les valeurs universelles des grandeurs qu'on cherche à estimer. Ces valeurs universelles ne peuvent être déterminées dans le cadre du système S de mesures dont on s'occupe, car par définition ce sont les moyennes d'un très grand nombre de déterminations obtenues à partir d'un très grand nombre de systèmes S.  
 (Dans les 4 premiers chapitres 1,2,3 et 4, nous supposons cependant que  $E(\dot{\mathcal{X}}) = \dot{\mathcal{X}}$ ).

${}_r\dot{\mathcal{X}}_1 = \dot{\mathcal{X}} - \mathcal{X}_0$

{ vecteur égal à  $\dot{\mathcal{X}}$  à une constante près  $\mathcal{X}_0$ , qui est une valeur approchée de  $\dot{\mathcal{X}}$ .  
 $\dot{\mathcal{X}}$  est donc un vecteur infiniment petit (sauf au cas où la liaison entre  $\dot{\mathcal{X}}$  et les grandeurs observées est linéaire, auquel cas on pourra prendre  $\mathcal{X}_0 = 0$  et  $\dot{\mathcal{X}} = \dot{\mathcal{X}}$ ).

$E({}_r\dot{\mathcal{X}}_1) + {}_r\mathcal{X}_{01}$

{ l'espérance mathématique de toutes les estimations des  $r$  grandeurs qu'on peut obtenir dans le cadre du système S de mesures.

${}_r\tilde{\mathcal{X}}_1$

{ une estimation de  $E(\dot{\mathcal{X}})$  obtenue dans le cadre du système S, par la méthode des moindres carrés : nous dirons aussi, en bref, une détermination de  $E(\dot{\mathcal{X}})$ .

${}_n\dot{L}_1 = \begin{pmatrix} \dot{L}_1 \\ \dot{L}_2 \\ \vdots \\ \dot{L}_n \end{pmatrix}$

{ les valeurs universelles des grandeurs observées.

$$\begin{aligned}
{}_n L_1 & \left\{ \begin{array}{l} \text{le vecteur-observation ; les observations peuvent être directes} \\ \text{et indépendantes, soit indirectes et corrélées.} \end{array} \right. \\
{}_n \tilde{L}_1 & \left\{ \begin{array}{l} \text{Le vecteur des estimations des grandeurs observées.} \\ \text{On dit classiquement, le vecteur des } \underline{\text{observations compensées}} \\ \text{On a : } E(\tilde{L}) = \dot{L} \end{array} \right. \\
{}_n e_1 = L - \dot{L} & \left\{ \begin{array}{l} \text{le vecteur-erreur de } L. \end{array} \right. \\
{}_n \dot{V}_1 = \dot{L} - L = -e & \left\{ \begin{array}{l} \text{le vecteur des } \underline{\text{résidus vrais}} : \text{ c'est l'opposé du vecteur-erreur.} \end{array} \right. \\
{}_n \tilde{V}_1 = \tilde{L} - L & \left\{ \begin{array}{l} \text{le vecteur des } \underline{\text{résidus}}, \text{ par définition : } \boxed{\text{résidu=compensé-observé}}. \end{array} \right. \\
\Gamma_L = P^{-1} \sigma_0^2 & \left\{ \begin{array}{l} \Gamma_L \text{ matrice variance des erreurs } e \text{ (exactitudes),} \\ P \text{ matrice des poids, évaluée à priori avant les calculs, symétrique,} \\ \text{définie positive non singulière, diagonale si les observations sont} \\ \text{indépendantes,} \\ \sigma_0^2 \text{ variance de la mesure de poids unitaire correspondant au choix de } P, \\ \text{si elle peut être fixée à priori, on appelle sa valeur } \underline{\text{variance unitaire à priori}}. \end{array} \right. \\
\text{Avec :} & \\
\Gamma_L = \Gamma_L^T & \\
\|\Gamma_L\| \neq 0 & \\
Z^T \Gamma_L Z > 0 \quad \forall Z \neq 0 & \\
L \in \mathcal{N}(\dot{L}, \sqrt{P^{-1}} \sigma_0) & \left\{ \begin{array}{l} \text{le vecteur } L \text{ est supposé (sauf contre-indication) normal,} \\ \text{sa densité de probabilité est :} \\ \frac{1}{\sqrt{(2\pi)^{n/2} |\text{Dét}(\Gamma_L)|}} e^{-\frac{1}{2\sigma_0^2} (L-\dot{L})^T P (L-\dot{L})} \end{array} \right.
\end{aligned}$$

## 2.2 POSE DES EQUATIONS

Les observations et les grandeurs à déterminer sont liées en général par des relations fonctionnelles  $\varphi(\dot{L}, \mathcal{X}) = 0$  plus ou moins compliquées et peu utilisables directement. Une des phases de l'analyse préliminaire consiste alors après avoir choisi les équations, à calculer ou à choisir des valeurs approchées des grandeurs à déterminer, valeurs au voisinage desquelles on développe les relations en se limitant au premier ordre.

On distingue traditionnellement deux manières de poser le problème de l'estimation de  $\mathcal{X}$  :

- par la pose d'équations d'observations,
- par la pose d'équations de condition.

### 2.2.1 Méthode des équations d'observations

Nous distinguons deux cas :

- celui où toutes les grandeurs à déterminer sont indépendantes (cas de l'exemple fondamental),
- celui où ces grandeurs sont liées par une ou plusieurs relations.

#### 2.2.1.1 Cas où toutes les grandeurs à déterminer sont indépendantes

Chaque grandeur observée  $L_i (i = 1, n)$  dépend de tout ou partie des grandeurs inconnues  $\mathcal{X}$  par l'intermédiaire d'une relation géométrique (ou physique) que nous supposons toujours de la forme :

$$\boxed{\varphi_i(L_i, \mathcal{X}) = 0 \quad (i = 1, n)} \quad (2.1)$$

(Si une liaison directe entre  $L_i$  et  $\mathcal{X}$  n'est pas possible, on est conduit à introduire des inconnues supplémentaires; cela se fait sans difficulté : ainsi si dans le cas de l'exemple fondamental on ne prenait qu'une seule inconnue, par exemple le côté  $c$  on serait conduit pour pouvoir exprimer les relations entre grandeurs observées et le côté  $c$ , à introduire deux autres inconnues supplémentaires, par exemple  $a$  et  $b$ ..

Pour qu'on n'ait pas à introduire d'inconnue supplémentaire il faut que soient réunies certaines conditions : par exemple que le nombre de degrés de liberté des  $r$  grandeurs inconnues soit égal à celui de la figure, autrement dit que les grandeurs inconnues définissent complètement la figure.

Le système (2.1) n'est pas linéaire en général, ni par rapport aux  $L_i$ , ni par rapport à  $\mathcal{X}$  ; or le problème que nous nous posons, à savoir déterminer la



meilleure estimation de  $\dot{\mathcal{X}}$  n'est soluble simplement que dans le cas où la liaison entre les  $\dot{L}_i$  et  $\dot{\mathcal{X}}$  est linéaire.

On se ramène à ce cas en linéarisant le système (2.1).

On calcule d'abord des valeurs approchées de toutes les inconnues, soit  $\mathcal{X}_0$  : cela est toujours possible puisque le nombre de grandeurs observées est suffisant pour que le problème soit géométriquement soluble.

On a donc :

$$\boxed{{}_r\dot{\mathcal{X}}_1 = {}_r\mathcal{X}_{01} + {}_r\dot{X}_1 \quad \text{avec} \quad {}_r\dot{X}_1 = \text{petit}} \quad (2.2)$$

Pour des raisons de symétrie et les besoins futurs du raisonnement nous écrirons aussi :

$$\dot{L}_i = L_{i,0} + (\dot{L}_i - L_{i,0}) \quad \text{avec} \quad L_{i,0} \text{ voisin de } \dot{L}_i \quad (2.3)$$

Le système (2.1) s'écrit alors :

$$\varphi_i(L_{i,0} + (\dot{L}_i - L_{i,0}), \mathcal{X}_0 + \dot{X}) = 0 \quad (i = 1, n)$$

Développons en série, et se limitant aux termes du premier ordre, on obtient le système linéaire suivant :

$$\varphi_i(L_{i,0}, \mathcal{X}_0) + \varphi'_{i(L_i)}(L_{i,0}, \mathcal{X}_0)(\dot{L}_i - L_{i,0}) - \alpha_{i1}\dot{X}_1 - \alpha_{i2}\dot{X}_2 - \dots - \alpha_{ir}\dot{X}_r = 0 \quad (i = 1, n)$$

qu'on peut encore écrire, en divisant par  $\varphi'_{i(L_i)}$  (différent de 0 sans quoi  $\dot{\mathcal{X}}$  ne dépendrait pas de  $\dot{L}_i$ ) :

$$\boxed{\begin{cases} a_{i1}\dot{X}_1 + a_{i2}\dot{X}_2 + \dots + a_{ir}\dot{X}_r + h_{i,0} + L_{i,0} - \dot{L}_i = 0 & (i = 1, n) \\ \text{Les constantes } a_{ij}, h_{i,0} \text{ ne dépendent que de } L_{i,0}, \text{ et } \mathcal{X}_0 \end{cases}} \quad (2.4)$$

Bien entendu un tel système n'est équivalent au système initial (2.1) que si les infiniments petits d'ordre supérieur à 1 sont numériquement négligeables, ce que nous supposons par la suite.

Il convient en particulier pour cela qu'on choisisse des valeurs approchées  $\mathcal{X}_0$  et  $L_0$  suffisamment proches de  $\dot{\mathcal{X}}$  et  $\dot{L}$ .

Ceci étant, si on pose :  $H_i = h_{i,0} + L_{i,0}$ , le système (2.4) s'écrit sous la forme :

$$\boxed{{}_n A_{r,r} \dot{X}_1 + {}_n H_1 = {}_n \dot{L}_1} \quad (2.5)$$

Un tel système linéaire ne fait que traduire à des infiniments petits près d'ordre 2 les liaisons géométriques entre grandeurs observées et grandeurs à déterminer.

Nous admettrons, dans tout ce qui suit qu'un tel système est équivalent au système initial  $\varphi_i(\dot{L}_i, \mathcal{X}) = 0$  ( $i = 1, n$ ) ce qui pratiquement est le cas dès qu'on possède des valeurs approchées  $\mathcal{X}_0$  suffisamment exactes <sup>a</sup>.

Et nous serons ramenés au problème suivant : estimer au mieux  $\dot{X}$  à partir d'un échantillon de  $n$  mesures  $L_i$  sachant que  $\dot{X}$  et  $\dot{L}$  sont liés linéairement selon (2.5).

a. Si tel n'est pas le cas, on serait conduit à réitérer le processus : étant calculée une première estimation  $\dot{X}$  on formera alors une nouvelle valeur approchée :  $\mathcal{X}_0 + \dot{X}$  et on procèdera à une nouvelle linéarisation, etc... En pratique le processus converge, et rapidement, même si les valeurs approchées de départ sont grossières. On s'arrête lorsque deux estimations successives ne diffèrent plus significativement.

Le système (2.5) peut encore s'écrire autrement ; tenant compte de :

$$\boxed{\begin{cases} \dot{L}_i = L_i - e_i = L_i + \dot{V}_i \\ (e_i \text{ erreur de } L_i; \dot{V}_i = \text{résidu vrai}) \end{cases}} \quad (2.6)$$

et posant :

$$K = H - L \quad (2.7)$$

il vient :

$$\boxed{\begin{cases} {}_n A_{r,r} \dot{X}_1 + {}_n H_1 = {}_n L_1 + {}_n \dot{V}_1 \\ A\dot{X} + K = \dot{V} \end{cases}} \quad (2.8)$$

qui est un système dit d'équations d'observations.

**Définition 2.1** Le système linéaire (impossible en général) à  $n$  équations et  $r$  inconnues :

$$\boxed{A.X + H = L \quad \text{ou} \quad A.X + K = o} \quad (2.9)$$

*s'appelle système des relations d'observations. L'élément  $AX + H$  s'appelle souvent élément calculé, et  $L$  l'élément observé.*

On peut montrer que la recherche du meilleur estimateur  $\tilde{X}$  de  $X$  au sens des moindres carrés, se ramène à la résolution par des moyens non classiques (inverse généralisée) d'un tel système insoluble généralement ( $n > r$ ) au sens classique, sauf dans deux cas : celui où  $n = r$  (pas de surabondance dans les observations), et celui où les observations sont sans erreurs (cas exclu).

Bien entendu la solution  $\tilde{X}$  que nous trouverons ne satisfera au système des relations d'observations que d'une manière approximative, d'autant meilleure que les erreurs de mesure seront petites.

On aura :

$$A\tilde{X} + H - L \neq 0$$

On posera alors :

$$\boxed{A\tilde{X} + H - L = \tilde{V} \quad \text{et} \quad \tilde{L} = L + \tilde{V}} \quad (2.10)$$

$\tilde{V}$  s'appelle le vecteur des résidus des relations d'observations et  $\tilde{L}$  le vecteur des observations compensées.

Par définition, on aura donc :

**Définition 2.2**

$$\boxed{A\tilde{X} + H = \tilde{L}} \quad (2.11)$$

soit, en remontant à la source :

$$\varphi_i(\mathcal{X}_0 + \tilde{X}, \tilde{L}_i) = 0; \quad (i = 1, n) \quad (2.12)$$

ensemble de relations exprimant que  $\mathcal{X}_0 + \tilde{X}$  et  $\tilde{L}$  sont les valeurs de grandeurs définissant une figure géométrique voisine de celle définie par  $\mathcal{X}$  et  $L$ .

Application à l'exemple fondamental :

Les observations :

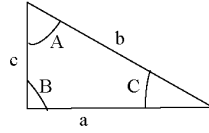


Fig. 2.1 L'Exemple fondamental

$$L = \begin{pmatrix} a \\ b \\ A \\ B \\ C \end{pmatrix}$$

Les grandeurs inconnues ont pour valeurs "vraies"  $\dot{a}, \dot{b}, \dot{c}$ ; les grandeurs observées ont pour valeurs vraies :  $\dot{a}, \dot{b}, \dot{A}, \dot{B}$  et  $\dot{C}$ ; les 5 relations mathématiques qui donnent les grandeurs observées en fonction des grandeurs inconnues sont (éléments observés ou deuxième membre, éléments calculés ou premier membre).

$$\begin{cases} \dot{a} = \dot{a} \\ \dot{b} = \dot{b} \\ \text{Arccos} \frac{\dot{b}^2 + \dot{c}^2 - \dot{a}^2}{2\dot{b}\dot{c}} = \dot{A} \\ \text{Arccos} \frac{\dot{c}^2 + \dot{a}^2 - \dot{b}^2}{2\dot{c}\dot{a}} = \dot{B} \\ \text{Arccos} \frac{\dot{a}^2 + \dot{b}^2 - \dot{c}^2}{2\dot{a}\dot{b}} = \dot{C} \end{cases} \quad (2.13)$$

On prendra comme valeurs approchées des inconnues :

$$a_0 = a; \quad b_0 = b; \quad c_0 = a \frac{\sin C}{\sin A}$$

Ceci étant, on posera :

- pour les grandeurs à déterminer :

$$\dot{a} = a_0 + da = a + da$$

$$\dot{b} = b_0 + db = b + db$$

$$\dot{c} = c_0 + dc$$

- pour les grandeurs observées :

$$\begin{aligned}\dot{a} &= a + \dot{V}_a; & \dot{A} &= A + \dot{V}_A; & \dot{C} &= C + \dot{V}_C \\ \dot{b} &= b + \dot{V}_b; & \dot{B} &= B + \dot{V}_B\end{aligned}$$

Il faut ensuite linéariser le système (2.13), c'est-à-dire les trois dernières relations :

- la première s'écrit :  $b^2 + c^2 - a^2 = 2\dot{b}\dot{c}\cos A$

soit :

$$b_0^2 + c_0^2 - a_0^2 + 2b_0db + 2c_0dc - 2a_0da = 2b_0c_0\cos A + 2c_0\cos Adb + 2b_0\cos Adc - 2b_0c_0\sin A \cdot \frac{\pi}{2000}\dot{V}_A$$

Posant alors :

$$k_A = \frac{b_0^2 + c_0^2 - a_0^2 - 2b_0c_0\cos A}{2b_0c_0\sin A}$$

on met sans difficulté la relation précédente sous la forme :

$$\frac{1}{\sin A} \frac{a_0}{b_0c_0} \frac{2000}{\pi} da - \frac{1}{\sin A} \frac{a_0^2 + b_0^2 - c_0^2}{2b_0^2c_0} \frac{2000}{\pi} db - \frac{1}{\sin A} \frac{a_0^2 + c_0^2 - b_0^2}{2b_0c_0^2} \frac{2000}{\pi} dc + k_A \frac{2000}{\pi} = \dot{V}_A$$

(étant entendu qu'on exprime  $\dot{V}_A$  en *dcgr*).

En définitive, on obtient le système suivant :

$$\begin{aligned}da &= \dot{V}_a \\ db &= \dot{V}_b \\ \frac{1}{\sin A} \frac{a_0}{b_0c_0} \frac{2000}{\pi} da - \frac{1}{\sin A} \frac{a_0^2 + b_0^2 - c_0^2}{2b_0^2c_0} \frac{2000}{\pi} db - \frac{1}{\sin A} \frac{a_0^2 + c_0^2 - b_0^2}{2b_0c_0^2} \frac{2000}{\pi} dc - k_A \frac{2000}{\pi} &= \dot{V}_A \\ -\frac{1}{\sin B} \frac{b_0^2 + a_0^2 - c_0^2}{2c_0a_0^2} \frac{2000}{\pi} da + \frac{1}{\sin B} \frac{b_0}{c_0a_0} \frac{2000}{\pi} db - \frac{1}{\sin B} \frac{b_0^2 + c_0^2 - a_0^2}{2a_0c_0^2} \frac{2000}{\pi} dc - k_B \frac{2000}{\pi} &= \dot{V}_B \\ \frac{1}{\sin C} \frac{c_0^2 + a_0^2 - b_0^2}{2a_0^2b_0} \frac{2000}{\pi} da - \frac{1}{\sin C} \frac{c_0^2 + b_0^2 - a_0^2}{2b_0^2a_0} \frac{2000}{\pi} db + \frac{1}{\sin C} \frac{c_0}{a_0b_0} \frac{2000}{\pi} dc - k_C \frac{2000}{\pi} &= \dot{V}_C\end{aligned}$$

Soit numériquement (unités normalisées : voir § 1.3) :

$$a_0 = 964.8 \quad b_0 = 1155.0 \quad c_0 = 633.6 \quad \text{unité : } 0.1 \text{ mm}$$

$$A = 630.42 \quad B = 998.02 \quad C = 370.08 \quad \text{unité : } 0.1 \text{ gr}$$

et :

$$\begin{pmatrix} 1. & 0. & 0. \\ 0. & 1. & 0. \\ 1.00375 & -0.83924 & 0.00143 \\ -1.00571 & 1.20285 & -0.66128 \\ 0.00094 & -0.36239 & 0.65918 \end{pmatrix} \cdot \begin{pmatrix} da \\ db \\ dc \end{pmatrix} - \begin{pmatrix} 0. \\ 0. \\ 0.97981 \\ -2.88449 \\ 0.42396 \end{pmatrix} = \begin{pmatrix} \check{V}_a \\ \check{V}_b \\ \check{V}_A \\ \check{V}_B \\ \check{V}_C \end{pmatrix} \quad (2.14)$$

On remarquera l'avantage d'utiliser des unités normalisées, c'est-à-dire des unités voisines de la valeur des  $emq$  : les quantités  $\check{V}_a, \dots, \check{V}_C$  (opposées aux erreurs) sont alors de l'ordre de l'unité, et il en est de même des quantités  $da, db, dc$  (si les valeurs approchées sont suffisamment bonnes) ; il en résulte que les coefficients des inconnues sont aussi de l'ordre de l'unité, si le système est géométriquement correctement conformé.

En choisissant d'autres unités, on peut au contraire faire apparaître de catégorie d'observations à catégorie d'observations des coefficients d'ordre de grandeur très différents, ce qui en cas de calcul manuel ou semi-manuel peut affaiblir considérablement la précision des calculs.

### 2.2.1.2 Cas où les $r$ grandeurs à déterminer sont liées par $p$ relations (méthode des équations d'observations : suite)

Nous supposons encore qu'il existe parmi les  $n$  grandeurs observées  $\check{L}$  un ensemble au moins, de  $r$  grandeurs observées permettant de déterminer les  $r$  grandeurs inconnues.

Autrement dit, nous nous bornerons à la considération du cas où la matrice  $A$  du système  $A\check{X} + H = \check{L}$  est comme précédemment de rang  $r$ .

Ceci étant, après (si nécessaire) linéarisation, les liaisons entre les  $r$  grandeurs inconnues, donneront lieu à un système de  $p$  équations de condition :

$$\begin{cases} {}_p B_{r,r} \check{X}_1 - M = 0 \\ \text{avec } B \text{ et } M \text{ ne dépendent que de } \mathcal{X} \\ \text{rang } B = p \end{cases} \quad (2.15)$$

Pour estimer au mieux  $\check{X}$  à partir de  $L$ , on aura donc outre l'échantillon  $L$  et sa loi de probabilité, à considérer l'ensemble des relations :

$$\begin{array}{l}
 {}_n A_{r,r} \dot{X}_1 + {}_n H_1 = {}_n \dot{L}_1 \\
 {}_p B_{r,r} \dot{X}_1 - {}_p M_1 = 0 \\
 \text{avec } \text{rang} A = r \\
 \text{rang} B = p
 \end{array}
 \tag{2.16}$$

(le cas où  $\text{rang} A = r$  et  $\text{rang} B = p$  est le plus simple à étudier, et probablement le plus intéressant en pratique).

Après compensation, on aura :

$$\begin{array}{l}
 A\tilde{X} + H = \tilde{L} \\
 B\tilde{X} - M = 0
 \end{array}$$

Faisons alors la remarque importante suivante : une équation de condition équivaut à une équation d'observation dont le poids serait infini : en effet  $M$  peut être considéré comme un vecteur d'observations fictives certaines, dont les  $emq$  sont nulles, et par suite les poids infinis. Une telle remarque est utile en pratique : au lieu d'utiliser la méthode correcte de "résolution" qui sera exposée un peu plus loin, on peut se borner à considérer qu'une équation de condition est une équation d'observations de poids très grand; on est ramené alors à traiter un système d'équations d'observations.

Il est clair, d'autre part, que ce cas se ramène aisément en principe au cas envisagé au § 2.2.1.1; il suffit en effet, à partir de  $B\tilde{X} - M = 0$  d'exprimer  $r$  inconnues en fonction des  $r - p$  autres, et de reporter les expressions correspondantes dans  $A\tilde{X} + H = L$ .

Néanmoins, il est bon d'étudier le cas sans tenir compte de cette possibilité : car dans certains problèmes on peut être amené à rajouter de telles équations de condition in extremis, ou bien à envisager les liaisons correspondantes à titre d'options. Il peut alors être commode de ne pas résoudre à priori le système  $B\tilde{X} - M = 0$ .

#### Exemple :

Reprenons le cas de l'exemple fondamental, mais en modifiant l'énoncé : nous nous proposerons de déterminer les mêmes inconnues  $\dot{a}, \dot{b}, \dot{c}$  à partir des mêmes mesures, mais en imposant au triangle d'être rectangle en  $B$  (cette condition pourrait être imposée en pratique pour diverses raisons parce que, par exemple, le triangle en question est un élément de maquette,

défini par des mesures directes et dont le projeteur désire à priori qu'il soit rectangle).

Pour obtenir l'équation de condition correspondante, on peut partir de la relation de Pythagore :

$$b^2 = a^2 + c^2$$

qui linéarisée donne :

$$a_0 da - b_0 db + c_0 dc + \frac{a_0^2 + c_0^2 - b_0^2}{2} = 0 \quad (2.17)$$

Soit numériquement (§ précédent) :

$$964.8 da - 1155.0 db + 633.6 dc - 868.5 = 0$$

ou encore :

$$da - 1.19714 db + 0.656722 dc - 0.90019 = 0 \quad (2.18)$$

### **2.2.2 Méthode des équations de condition liant les $n$ grandeurs observées ( $r$ grandeurs inconnues indépendantes; $r =$ nombre de degrés de liberté de la figure ou du phénomène)**

Si  $r$  est le nombre de degrés de liberté de la figure, c'est qu'il y a  $(n - r)$  relations indépendantes entre les  $n$  grandeurs observées :

$$\boxed{\begin{cases} \Psi_i(\hat{L}) = 0 \\ (i = 1, n - r) \end{cases}} \quad (2.19)$$

Un tel système exprime aussi bien que celui des équations d'observations, la géométrie d'une figure dépendant de  $r$  paramètres; c'est un système d'équations de condition liant les grandeurs observées.

Il donne lieu à la méthode dite des équations de condition liant les grandeurs observées, variante importante en pratique de la méthode des équations d'observations.



Le système (2.19) doit être linéarisé pour être utilisable; nous prendrons comme valeurs approchées de  $\dot{L}$ , les valeurs observées  $L$  :

$$\dot{L} = L + \dot{V} \quad (\dot{V} \text{ résidus vrais})$$

D'où, en développant et se limitant aux termes du premier ordre :

$$\begin{cases} \Psi_i(L) + C_i \dot{V} = 0 \\ (i = 1, n-r) \end{cases}$$

Soit :

$$\boxed{\begin{array}{l} {}_{n-r}C_{n-n} \dot{V}_1 - {}_{n-r}Q_1 = 0 \\ \text{rang} C = n-r \end{array}} \quad (2.20)$$

Ce système linéaire est équivalent au système (2.19) dans la mesure où les termes négligés sont effectivement négligeables; on remarquera qu'il y a ici moins de précautions à prendre que dans le cas des équations d'observations, du fait que les valeurs approchées ici, sont égales aux valeurs observées, donc très proches des valeurs vraies  $\dot{L}$  (les termes négligés sont donc de l'ordre du carré des erreurs d'observations, l'infiniment petit principal étant une grandeur observée moyenne).

Par contre surgit en pratique une difficulté : celle de trouver  $(n-r)$  relations indépendantes liant les  $n$  grandeurs observées, dans le cas où le réseau est un peu complexe; on est sûr de leur existence mais le problème peut être de les trouver.

Faisons trois remarques importantes qui nous serviront par la suite :

**Remarque 2.1** *Le système d'équations de condition est défini à un facteur multiplicatif matriciel régulier près.*

En effet étant donné n'importe quelle matrice carrée régulière d'ordre  $(n-r)$ ,  $D$ , on a :

$$\boxed{C\dot{V} - Q = 0} \iff \boxed{\begin{array}{l} D.C.\dot{V} - D.Q = 0 \\ |D| \neq 0 \end{array}} \quad (2.21)$$

**Remarque 2.2** *Le système d'équations de condition est formellement équivalent à un système de  $n$  équations d'observations à  $n$  inconnues liées par  $(n-r)$  équations de condition.*

En effet :

$$\boxed{C\dot{V} - Q = 0} \iff \boxed{\begin{array}{l} \dot{Y} = \dot{V} \\ C\dot{Y} - Q = 0 \end{array}} \quad (2.22)$$

**Remarque 2.3** *A des infiniments petits près du 2ème ordre (carré des erreurs) le système des équations d'observations  $A\dot{X} - K = \dot{V}$  ( $r$  inconnues indépendantes) et celui des équations de condition se déduisent l'un de l'autre par des substitutions linéaires.*

En effet :

\* soit d'abord le système des équations d'observations :  $A\dot{X} - K = \dot{V}$  ; puisque le problème est soluble géométriquement,  $A$  est de rang  $r$  on peut donc calculer  $\dot{X}$  en fonction de  $r$  résidus vrais que nous pourrons toujours supposer être les premiers :  $\dot{V}_1, \dot{V}_2, \dots, \dot{V}_r$

Ce qui revient à dire que parmi les  $n$  grandeurs observées, il en existe  $r$  en fonction desquelles on peut calculer les  $r$  inconnues :

$${}_r\dot{X}_1 = {}_rB_r \begin{pmatrix} \dot{V}_1 \\ \dot{V}_2 \\ \vdots \\ \dot{V}_r \end{pmatrix} + {}_rD_1 \quad (2.23)$$

Reportant alors dans les  $(n - r)$  équations restantes, on obtient :

$$\begin{pmatrix} \dot{V}_{r+1} \\ \dot{V}_{r+2} \\ \vdots \\ \dot{V}_n \end{pmatrix} = E \begin{pmatrix} \dot{V}_1 \\ \dot{V}_2 \\ \vdots \\ \dot{V}_r \end{pmatrix} + {}_{n-r}FD_1 \quad (2.24)$$

ensemble de relations linéaires exprimant que les  $(n - r)$  résidus vrais relatifs aux  $(n - r)$  autres grandeurs observées sont des fonctions linéaires des  $r$  premiers ou encore que les  $(n - r)$  grandeurs observées  $\dot{L}_{r+1}, \dots, \dot{L}_n$  sont des fonctions des  $r$  premières  $\dot{L}_1, \dot{L}_2, \dots, \dot{L}_r$ .

Le système (2.24) se met sous la forme :

$$\begin{array}{l} C'\dot{V} - Q' = 0 \\ \text{avec rang } C = n - r \end{array} \quad (2.25)$$

Ce système aussi bien que (2.20),  $C\dot{V} + Q = 0$  exprime au 2ème ordre près que le nombre de degrés de liberté des  $n$  grandeurs observées est  $r$ ; il s'ensuit que leurs premiers membres sont identiques à un facteur matriciel régulier près :

$$\exists {}_{n-r}D_{n-r} \text{ tel que } {}_{n-r}C'_r = {}_{n-r}D_{n-r}C$$

$${}_{n-r}Q'_r = {}_{n-r}D_{n-r}Q$$

avec  $D$  régulière, soit :  $C'\dot{V} + Q' = 0 \iff C\dot{V} + Q = 0$ .

**Conclusion 2.1** *La substitution linéaire :*

$${}_r\dot{X}_1 = {}_rB_r \begin{pmatrix} \dot{V}_1 \\ \dot{V}_2 \\ \vdots \\ \dot{V}_r \end{pmatrix} + D = {}_rB'_n \dot{V} + D$$

*transforme le système des équations d'observations  $A\dot{X} - K = \dot{V}$  en le système d'équations de condition  $C\dot{V} - Q = 0$  (défini à un facteur matriciel régulier d'ordre  $(n - r)$  près).*

\* Réciproquement partons du système d'équations d'observations :

$$C\dot{V} - Q = 0$$

Ce système est équivalent à :

$$\begin{aligned} \dot{Y} &= \dot{V} \\ C\dot{Y} - Q &= 0 \end{aligned} \quad (2.26)$$

Faisons alors la substitution  $\dot{Y} = A\dot{X} - K$ ; on obtient :

$$\begin{aligned} A\dot{X} - K &= \dot{V} \\ CA\dot{X} - CK - Q &= 0 \end{aligned}$$

La deuxième équation exprime que les grandeurs inconnues sont liées, ce qui est contraire aux hypothèses; par suite :

$$CA = 0, \quad CK + Q = 0 \quad (2.27)$$

Par suite, il subsiste seulement :

$$A\dot{X} - K = \dot{V}$$

**Conclusion 2.2** La substitution linéaire  $\dot{Y} = A\dot{X} - K$  transforme le système des équations de condition écrit sous la forme :

$$\dot{Y} = \dot{V}, \quad C\dot{Y} - Q = 0$$

en le système des équations d'observations  $A\dot{X} - K = \dot{V}$  (à des termes du 2ème ordre près).

### Application à l'exemple fondamental

Les grandeurs observées sont ici,  $\dot{a}, \dot{b}, \dot{A}, \dot{B}$  et  $\dot{C}$  pour trois inconnues  $\dot{a}, \dot{b}$  et  $\dot{c}$ .

Il y a donc deux équations de condition à écrire ; nous prendrons :

$$\dot{A} + \dot{B} + \dot{C} = \pi$$

$$\frac{\sin \dot{A}}{\dot{a}} = \frac{\sin \dot{B}}{\dot{b}}$$

soit encore en exprimant les grandeurs observées en fonction de leurs résidus vrais :

$$\begin{aligned} \dot{V}_A + \dot{V}_B + \dot{V}_C + A + B + C - \pi &= 0 \\ \frac{\sin A}{a} + \frac{\cos A}{a} \dot{V}_A - \frac{\sin A}{a^2} \dot{V}_a &= \frac{\sin B}{b} + \frac{\cos B}{b} \dot{V}_b - \frac{\sin B}{b^2} \dot{V}_b \end{aligned}$$

Soit :

$$\begin{aligned} \dot{V}_A + \dot{V}_B + \dot{V}_C + A + B + C - \pi &= 0 \\ -\sin A \dot{V}_a + \frac{a^2}{b^2} \sin B \dot{V}_b + a \cos A \dot{V}_A - \frac{a^2}{b} \cos B \dot{V}_B + a \sin A - \frac{a^2}{b} \sin B &= 0 \end{aligned} \quad (2.28)$$

### Application numérique :

En unités normalisées (c'est-à-dire proches de la valeur de l'écart-type de chaque catégorie d'observations), les mesures sont :

$$\begin{aligned} a &= 964.8 \quad \text{unité normalisée : le } dcmm \\ b &= 1150.0 \end{aligned}$$

$$A = 964.80$$

$$B = 998.02 \quad \text{unité normalisée : le } dcgr$$

$$C = 370.08$$

Tenant compte de ces unités (2.28) s'écrit matriciellement :

$$\begin{pmatrix} 0.00 & 0.00 & 1.00 & 1.00 & 1.00 \\ -0.83617 & 0.69776 & 0.83121 & 0.00394 & 0.00 \end{pmatrix} \begin{pmatrix} \dot{V}_a \\ \dot{V}_b \\ \dot{V}_A \\ \dot{V}_B \\ \dot{V}_C \end{pmatrix} + \begin{pmatrix} -1.48000 \\ 0.81888 \end{pmatrix} = 0 \quad (2.29)$$

On voit que là aussi l'emploi des unités normalisées conduit mécaniquement à des coefficients bien cadrés; à ceci près que l'on n'oubliera pas qu'il est licite de multiplier par n'importe quel coefficient une équation de condition (ceci est faux pour une relation d'observations).

Remarque : Nous avons noté qu'étant donné le système des équations d'observations  $A\dot{X} - K = \dot{V}$ , et le système d'équations de condition  $C\dot{V} - Q = 0$ , on devrait avoir (2.27) :

$$C.A = 0, \quad \text{et} \quad C.K + Q = 0$$

Si on essaie de vérifier ceci pour l'exemple fondamental on trouve que cela est vrai à 0.001 près seulement, et non pas à 0.00001 près (précision des calculs numériques), comme on pourrait s'y attendre.

Ceci est dû au fait que de telles égalités ne sont valables qu'au 2ème ordre près, puisque dans les calculs de la linéarisation qui conduisent aux systèmes correspondants on néglige systématiquement les infiniments petits d'ordre supérieur.

Donnons un exemple : dans (2.28),  $(\sin\dot{A})V_a$  a été assimilé à  $(\sin A)\dot{V}_a$  : on a ainsi commis une erreur de :

$$(\sin\dot{A})V_a - (\sin A)\dot{V}_a \approx (\cos A)\dot{V}_A\dot{V}_a$$

or  $\dot{V}_A\dot{V}_a$  est de l'ordre de 0.001...

On voit qu'il suffirait en principe d'effectuer ici les calculs avec quatre chiffres derrière la virgule.

### 2.3 DÉTERMINATION D'UN ESTIMATEUR $\tilde{X}$ DE $\dot{X}$ PAR LA MÉTHODE DES MOINDRES CARRÉS - ESTIMATEUR D'UNE FONCTION LINÉAIRE DE $\dot{X}$ - ESTIMATEURS DE $\dot{L}$ ET $\dot{V}$ - INVARIANCE DE LA SOLUTION PAR CHANGEMENT D'INCONNUES

#### 2.3.1 Problème linéaire

Nous avons ramené le problème de l'estimation de  $r$  grandeurs inconnues  ${}_r\dot{X}_1$  ; à un problème linéaire, de 2 façons :

\* par la méthode des équations d'observations qui consiste à exprimer les  $n$  grandeurs observées  ${}_n\dot{L}_1$  en fonction des  $r$  grandeurs inconnues  ${}_r\dot{X}_1$ , celles-ci étant exprimées sous la forme  $\dot{\mathcal{X}} = \mathcal{X}_0 + \dot{X}$  ( $\mathcal{X}_0$  valeurs approchées de  $\dot{\mathcal{X}}$ ). On obtient ainsi un système linéaire de la forme :

$${}_nA_{r,r}\dot{X}_1 - {}_nK_1 = {}_n\dot{V}_1 = \dot{L} - L$$

avec  $\text{rang}A = r$ , c'est-à-dire nous supposons qu'il existe au moins un ensemble de  $r$  grandeurs observées qui déterminent géométriquement  $\dot{X}$ .

Nous avons supposé de plus que les  $r$  grandeurs inconnues pouvaient être liées par  $p$  relations linéaires indépendantes :

$${}_pB_{r,r}\dot{X}_1 - {}_pM_1 = 0 \text{ avec } \text{rang}B = p$$

En résumé, le problème de l'estimation de  $\dot{X}$  (et donc de  $\dot{\mathcal{X}}$ ) se pose mathématiquement sous la forme suivante :

$$\left\{ \begin{array}{l} {}_nA_{r,r}\dot{X}_1 - {}_nK_1 = {}_n\dot{V}_1 = \dot{L} - L \\ {}_pB_{r,r}\dot{X}_1 - {}_pM_1 = 0 \\ \text{rang}A = r, \text{ rang}B = p \\ L \in \mathcal{N}(\dot{L}, \sqrt{P^{-1}}\sigma_0) \end{array} \right. \quad (2.30)$$

\* par la méthode des équations de condition liant les grandeurs observées  ${}_n\dot{L}_i$  (nous avons supposé que dans ce cas, les  $r$  grandeurs inconnues étaient indépendantes).

$${}_{n-r}C_{nn}\dot{V}_1 - {}_{n-r}Q_1 = 0, \quad \text{avec } \text{rang } C = n - r$$

système qui, comme nous l'avons remarqué, peut s'écrire aussi :

$$\boxed{\begin{cases} {}_n\dot{Y}_1 = {}_n\dot{V}_1 \\ {}_{n-r}C_{nn}\dot{Y}_1 - {}_{n-r}Q_1 = 0 \\ \text{rang } C = n - r \end{cases}} \quad (2.31)$$

et qui est de la forme (2.30) avec  $A = I$ , de  $\text{rang } n$  égal au nombre d'inconnues à déterminer, ces inconnues étant liées par  $(n - r)$  relations indépendantes.

L'estimation de  $\dot{Y}$ , c'est-à-dire de  $\dot{V}$ , dans le cas de pose d'équations de condition, se traitera donc par la même méthode.

Une fois  $\dot{V}$  estimé, on estimera  $\dot{X}$  ( $\implies \dot{\mathcal{X}} = \dot{\mathcal{X}}_0 + \dot{X}$ ) à l'aide d'un ensemble de relations linéaires (qu'il conviendra d'établir) de la forme  $\dot{X} = U\dot{V} + W$ .

Bien entendu; il n'est pas évident a priori, que les 2 méthodes conduisent à des estimations identiques de  $\dot{\mathcal{X}}$ ; il convient de la démontrer; nous verrons d'ailleurs que ces estimations ne sont les mêmes qu'au deuxième ordre près.

### 2.3.2 Détermination de l'estimateur $\tilde{X}$ de $\dot{X}$

Il nous faut donc à présent résoudre le problème suivant; estimer au mieux  $\dot{X}$  sachant que :

$$\left\{ \begin{array}{l} {}_nA_{rr}\dot{X}_1 - {}_nK_1 = {}_n\dot{L}_1 - {}_nL_1 \\ B\dot{X} - M = 0 \\ \text{rang } A = r, \text{ rang } B = p \\ L \in \mathcal{N}(\dot{L}, \sqrt{P^{-1}}\sigma_0) \end{array} \right. \quad (2.32)$$

**Théorème 2.1** *En minimisant la forme quadratique fondamentale :*

$$\boxed{V^T P V = (A\dot{X} - K)^T P (A\dot{X} - K)} \quad (2.33)$$

compte tenu de la condition  $BX - M = 0$ , on obtient un estimateur de  $\dot{X}$  asymptotiquement sans biais, asymptotiquement efficace et asymptotiquement normal.<sup>1</sup>

En effet, le vecteur aléatoire  $L$ , obéissant à la loi normale  $\mathcal{N}(\dot{L}, \sqrt{P^{-1}}\sigma_0)$ , sa densité de probabilité  $f(L)$  est :

$$f(L) = \frac{1}{(2\pi)^{n/2}} \frac{\sqrt{|P|}}{\sigma_0^{2n/2}} e^{-\frac{1}{2\sigma_0^2} (L - \dot{L})^T P (L - \dot{L})} \quad (2.34)$$

soit encore :

$$f(X) = \frac{1}{(2\pi)^{n/2}} \frac{\sqrt{|P|}}{(\sigma_0^2)^{n/2}} e^{-\frac{1}{2\sigma_0^2} (A\dot{X} - K)^T P (A\dot{X} - K)} \quad (2.35)$$

Sous cette forme il est clair que  $f$  ne dépend que des  $(r + 1)$  paramètres  ${}_r\dot{X}_1, \sigma_0^2$  liés eux mêmes par  ${}_pB_r{}_r\dot{X}_1 - M = 0$ , donc en fait de  $(r + 1 - p)$  paramètres.

La règle du maximum de vraisemblance conduit à adopter comme estimateurs  $\tilde{X}$  et  $s_0^2$  de  $\dot{X}$  et  $\sigma_0^2$ , ceux qui rendent maximum  $f$  ou encore  $\text{Log}f$  compte tenu des équations de condition  $B\tilde{X} - M = 0$ .

On doit donc avoir :

$$\left[ -\frac{n}{2} \text{Log} s_0^2 - \frac{1}{2s_0^2} (A\tilde{X} - K)^T P (A\tilde{X} - K) \right] \text{ maximal} \quad (2.36)$$

$$B\tilde{X} - M = 0$$

Pour résoudre ce problème on peut employer la méthode des "extremums liés" (méthode des multiplicateurs de Lagrange) qui consiste à considérer la fonction scalaire :

$$U = -\frac{n}{2} \text{Log} s_0^2 - \frac{1}{2s_0^2} (A\tilde{X} - K)^T P (A\tilde{X} - K) - \lambda^T (B\tilde{X} - M)$$

( $\lambda$  étant un vecteur inconnu à  $p$  composantes) et à exprimer que les dérivées de  $U$  par rapport à  $\tilde{X}, s_0^2$  et  $\lambda$  sont nulles;  $\tilde{X}, s_0^2$  et  $\lambda$  sont alors solutions du

1. On montrera plus tard (chapitre 4) qu'en fait l'estimateur  $\tilde{X}$  est sans biais (au sens statistique), totalement efficace, et normal.



système ainsi obtenu, soit :

$$\begin{aligned}\frac{\partial U}{\partial \tilde{X}} &= -\frac{1}{2s_0^2} \frac{\partial}{\partial \tilde{X}} (\tilde{X}^T A^T P A \tilde{X} - 2\tilde{X}^T A^T P K + K^T P K) - B^T \lambda = 0 \\ \frac{\partial U}{\partial s_0^2} &= -\frac{n}{2s_0^2} + \frac{1}{2s_0^4} (A\tilde{X} - K)^T P (A\tilde{X} - K) = 0 \\ \frac{\partial U}{\partial \lambda} &= -(B\tilde{X} - M) = 0\end{aligned}$$

qui s'écrit encore en posant  $\Lambda = \lambda s_0^2$  :

$$\begin{aligned}A^T P A \tilde{X} + B^T \Lambda - A^T P K &= 0 \\ B\tilde{X} - M &= 0 \\ s_0^2 &= \frac{1}{n} (A\tilde{X} - K)^T P (A\tilde{X} - K)\end{aligned}$$

soit encore :

$$\boxed{\begin{pmatrix} {}_r A_n^T P_n A_r & {}_r B_p^T \\ {}_p B_r & 0 \end{pmatrix} \cdot \begin{pmatrix} \tilde{X} \\ \Lambda \end{pmatrix} = \begin{pmatrix} A^T P K \\ M \end{pmatrix}} \quad (2.37)$$

$$s_0^2 = \frac{1}{n} (A\tilde{X} - K)^T P (A\tilde{X} - K)$$

D'après la théorie de la règle du maximum de vraisemblable (cours statistique) on peut alors affirmer :

- que la solution  $\tilde{X}$ ,  $s_0^2$  existe,
- qu'elle maximise la fonction de vraisemblable  $f$ ; on voit immédiatement en outre qu'elle minimise la forme quadratique  $(AX - K)^T P (AX - K)$ ,
- que  $\tilde{X}$  et  $s_0^2$  sont des estimateurs asymptotiquement sans biais, efficaces et normaux.

On peut établir directement certains de ces résultats, en étudiant les caractéristiques du système (2.37), compte tenu des hypothèses (2.32); nous allons le faire, car cela nous permettra d'une part d'établir un certain nombre de propriétés utiles, et d'autre part de généraliser la notion d'estimateur des moindres carrés au cas où  $L$  n'est plus un vecteur - observation normal.

**Théorème 2.2** *La solution du système :*

$$\begin{pmatrix} {}_r A_n^T P_n A_r & {}_r B_p^T \\ {}_p B_r & 0 \end{pmatrix} \cdot \begin{pmatrix} {}_r \tilde{X}_1 \\ {}_p \Lambda_1 \end{pmatrix} = \begin{pmatrix} {}_r A_n^T P_n K_1 \\ {}_p M_1 \end{pmatrix} \quad (2.38)$$

compte tenu des seules hypothèses  $\text{rang} A = r$ , et  $\text{rang} B = p$ ,  $P$  régulière et symétrique définie positive; elle est unique et minimise la forme quadratique fondamentale :

$$V^T P V = (A X - K)^T P (A X - K)$$

Nous dirons qu'elle minimise la somme des carrés des résidus pondérés; (la solution  $\tilde{X}$  des moindres carrés apparaît donc de ce fait comme la généralisation de l'estimateur moyenne expérimentale dans le cas de  $n$  mesures).

\*\* Démontrons d'abord que le système (2.38) est possible; on peut l'écrire :

$$\begin{aligned} A^T P A \tilde{X} + B^T \Lambda - A^T P K &= 0 \\ B \tilde{X} - M &= 0 \end{aligned} \quad (2.39)$$

avec  $\text{rang} A = r$  et  $\text{rang} B = n - r$

Etudions d'abord la matrice :

$${}_r N_r = A^T P A = (\sqrt{P} A)^T (\sqrt{P} A) \quad (2.40)$$

qu'on appelle la matrice normale.

\* La matrice normale est symétrique, définie positive (donc régulière) : il est évident que  $N = N^T$ ; d'autre part le fait que  $A$  est de rang  $r$ , et  $P$  de rang  $n > r$  implique que  ${}_r N_r$  est aussi de rang  $r$  (car  $PA$  est de rang  $r$  et  $A^T P A$  par suite également), et donc inversible; enfin  $N$  est définie positive, on entend par là que :

$$x^T N x > 0, \forall x \neq 0$$

Cela est évident :

$$x^T N x = (\sqrt{P} A x)^T (\sqrt{P} A x) > 0, \forall x \neq 0$$

(Il revient au même de dire que toutes les valeurs propres de  $N$  sont positives).

\* L'inverse de la matrice normale  $N^{-1}$  est également symétrique définie positive. C'est évident.

Ceci étant, puisque  $N^{-1}$  existe, on a de (2.39) :

$$\tilde{X} = N^{-1} (A^T P K - B^T \Lambda) \quad (2.41)$$

puis en reportant dans la deuxième équation de (2.39), on obtient :

$$B N^{-1} B^T \Lambda = B N^{-1} A^T P K - M \quad (2.42)$$

Il est aisé de voir que la matrice  $BN^{-1}B^T$  tout comme la matrice normale  $N$ , est symétrique définie positive, donc inversible; en effet  $N^{-1}$  est symétrique, et peut donc s'écrire sous la forme  $\sqrt{N^{-1}}^T \cdot \sqrt{N^{-1}}$ ;  $B$  étant de rang  $p$  et  $\sqrt{N^{-1}}$  de rang  $r > p$  le produit  $B\sqrt{N^{-1}}$  est de rang  $p$  et le produit  $B\sqrt{N^{-1}}$  par sa transposée est encore de rang  $p$ ;  $BN^{-1}B^T$  est donc régulière; on voit sans peine qu'elle est aussi symétrique définie positive.

D'où il suit (2.42) :

$$\Lambda = (BN^{-1}B^T)^{-1}(BN^{-1}A^T PK - M)$$

et :

$$\boxed{\begin{array}{l} \tilde{X} = N^{-1} [I - B^T (BN^{-1}B^T)^{-1} BN^{-1}] A^T PK + N^{-1} B^T (BN^{-1}B^T)^{-1} M \\ \text{avec } N = A^T PA \end{array}} \quad (2.43)$$

\*\* La solution  $\tilde{X}$  minimise la somme des carrés des résidus pondérés  $V^T PV$  : elle est bien la solution des "moindres carrés".

En effet, posant  $\tilde{V} = A\tilde{X} - K$ , on a :

$$\begin{array}{l} A\tilde{X} - K = \tilde{V} \\ B\tilde{X} - M = 0 \\ A^T PA\tilde{X} + B^T \Lambda = A^T PK \end{array} \quad (2.44)$$

(puisque  $\tilde{X}$  satisfait à (2.39)).

Considérons alors un autre estimateur de  $\hat{X}$ , soit  $X \neq \tilde{X}$  satisfaisant lui aussi à  $BX - M = 0$ , et posant :

$$AX - K = V$$

On a donc :

$$\begin{array}{l} AX - K = V \\ BX - M = 0 \end{array} \quad (2.45)$$

Nous allons montrer que :  $V^T PV > \tilde{V}^T P\tilde{V}$ .

En effet, on voit d'abord que :

$$A\tilde{X} - K = \tilde{V} \text{ et } AX - K = V \implies V - \tilde{V} = A(X - \tilde{X})$$

On peut donc écrire :

$$V^T PV = (\tilde{V} + A(X - \tilde{X}))^T P (\tilde{V} + A(X - \tilde{X}))$$

ou :

$$V^T PV = \tilde{V}^T P \tilde{V} + (X - \tilde{X})^T A^T P A (X - \tilde{X}) + 2(X - \tilde{X})^T A^T P \tilde{V} \quad (2.46)$$

Mais de (2.44), on déduit (première relation multipliée par  $A^T P$ , et troisième relation) :

$$\begin{aligned} A^T P A \tilde{X} - A^T P K &= A^T P \tilde{V} \\ A^T P A \tilde{X} + B^T \Lambda - A^T P K &= 0 \end{aligned}$$

ce qui entraîne :

$$A^T P \tilde{V} = -B^T \Lambda \quad (2.47)$$

et permet d'écrire :

$$(X - \tilde{X})^T A^T P \tilde{V} = -(\Lambda^T B (X - \tilde{X}))^T$$

Mais par hypothèse :

$$B\tilde{X} - M = 0 \text{ et } BX - M = 0 \implies B(X - \tilde{X}) = 0$$

et par suite :

$$(X - \tilde{X})^T A^T P \tilde{V} = 0 \quad (2.48)$$

Soit en définitive, en revenant à (2.46) :

$$V^T PV = \tilde{V}^T P \tilde{V} + (\sqrt{P}A(X - \tilde{X}))^T (\sqrt{P}A(X - \tilde{X})) > \tilde{V}^T P \tilde{V}$$

**Conclusion 2.3** De toutes les solutions  $X$  qui respectent la géométrie de la figure, celle des moindres carrés, est la seule qui minimise la somme des carrés des résidus pondérés  $V^T PV$  (d'où son nom).

Notons en passant la formule (2.47) :

$$\boxed{A^T P \tilde{V} = -B^T \Lambda} \quad (2.49)$$

qui s'écrit quand  $B = 0$  :

$$A^T P \tilde{V} = 0 \quad (2.50)$$

Généralisation d'une propriété classique de la moyenne de  $n$  nombres : la somme des écarts algébriques à la moyenne, c'est-à-dire des résidus, est nulle.

Ainsi la formule :

$$\tilde{V}^T P \tilde{V} = K^T P K - 2B^T \Lambda - \tilde{X}^T N \tilde{X} \quad (2.51)$$

qu'on établit sans difficulté, et qui peut être utile pour estimer la variance unitaire lorsqu'on a perdu de mémoire les relations d'observations (Nous verrons que la variance unitaire peut être estimée par  $\frac{\tilde{V}^T P \tilde{V}}{n-r}$ ).

### 2.3.3 Estimateur d'une fonction linéaire de $\dot{X}$

**Théorème 2.3** *L'estimateur au sens des moindres carrés d'une fonction de  $\dot{X}$ ,  $\dot{Y} = D\dot{X} + E$  est  $\tilde{Y} = D\tilde{X} + E$ ,  $\tilde{X}$  étant l'estimateur de  $\dot{X}$  au sens des moindres carrés.*

Ce résultat est normal, mais n'est pas évident : car le problème est à présent d'estimer  $\dot{Y}$  à partir de l'échantillon  $L$ , et non plus seulement  $\dot{X}$ , et rien ne permet d'affirmer a priori que la solution est celle de l'énoncé du théorème.

Comme hypothèse, nous avons :

$$\begin{cases} {}_n A_r \dot{X}_1 - K = \dot{L} - L & \text{avec } \text{rang} A = r \\ {}_p B_r \dot{X}_1 - M = 0 & \text{avec } \text{rang} B = p \\ {}_s \dot{Y}_1 = {}_s D_r \dot{X}_1 + E \\ L \in \mathcal{N}(\dot{L}, \sqrt{P^{-1}} \sigma_0) \end{cases} \quad (2.52)$$

Il s'agit donc cette fois, si on pose :

$$\dot{Z} = \begin{pmatrix} {}_r \dot{X}_1 \\ {}_s \dot{Y}_1 \end{pmatrix}$$

d'estimer le vecteur  $\dot{Z}$ , sachant que :

$$(A, 0) \cdot \tilde{Z} - K = \hat{V}$$

$$\begin{pmatrix} B & 0 \\ D & -I \end{pmatrix} \cdot \tilde{Z} - \begin{pmatrix} M \\ -E \end{pmatrix} = 0 \quad (2.53)$$

En suivant la même démarche qu'au § 2.3.2, on est alors conduit à minimiser la forme quadratique fondamentale, compte tenu des relations entre les  $(r + s)$  inconnues; c'est-à-dire à considérer,  $\tilde{Z}$  étant l'estimateur cherché, et  $\Lambda$  un certain vecteur constant inconnu, la fonction scalaire :

$$U = ((A, 0)\tilde{Z} - K)^T P((A, 0)\tilde{Z} - K) - 2\Lambda^T \left( \begin{pmatrix} B & 0 \\ D & -I \end{pmatrix} \cdot \tilde{Z} - \begin{pmatrix} M \\ -E \end{pmatrix} \right)$$

et à exprimer que ses dérivées partielles par rapport à  $\tilde{Z}$  et  $\Lambda$  sont nulles :

$$\frac{\partial U}{\partial \tilde{Z}} = \begin{pmatrix} N & 0 \\ 0 & 0 \end{pmatrix} \cdot \tilde{Z} - \begin{pmatrix} A^T P K \\ 0 \end{pmatrix} - \begin{pmatrix} B^T & D^T \\ 0 & -I \end{pmatrix} \Lambda = 0$$

$$\frac{\partial U}{\partial \Lambda} = \begin{pmatrix} B & 0 \\ D & -I \end{pmatrix} \cdot \tilde{Z} - \begin{pmatrix} M \\ -E \end{pmatrix}$$

Soit encore, en revenant à  $\tilde{X}$  et  $\tilde{Y}$  et en posant :

$$\Lambda = \begin{pmatrix} p\Lambda_{11} \\ s\Lambda_{21} \end{pmatrix}$$

ou encore :

$$\begin{aligned} N\tilde{X} - B^T \Lambda - A^T P K &= 0 \\ B\tilde{X} - M &= 0 \\ \tilde{Y} &= D\tilde{X} + E \end{aligned} \quad (2.54)$$

Les deux premières équations redonnent pour  $\tilde{X}$ , le même estimateur que celui déjà trouvé au § 2.3.2 et l'estimateur au sens des moindres carrés de  $\tilde{Y}$  est bien  $D\tilde{X} + E$ .

### 2.3.4 Estimateurs de $\hat{L}$ et $\hat{V}$

De  $\hat{V} = A\tilde{X} - K$  et  $\hat{L} = L + \hat{V}$  et du théorème précédent, on déduit que les estimateurs au sens des moindres carrés de  $\hat{V}$  et  $\hat{L}$  sont respectivement :

$$\boxed{\hat{V} = A\tilde{X} - K, \quad \hat{L} = L + \hat{V}} \quad (2.55)$$

### 2.3.5 Invariance de la solution par changement d'inconnues

Dans le cas de l'exemple fondamental, nous avons choisi comme inconnues pour déterminer le triangle ses côtés :  $\dot{a}, \dot{b}, \dot{c}$  mais nous aurions pu en choisir d'autres par exemple l'aire, le périmètre, une hauteur et un côté (soit 4 inconnues, non indépendantes).

La méthode des moindres carrés permet d'estimer directement ces 4 inconnues à partir de mesures  $L$ ; ces estimations sont-elles identiques à celles qu'on peut calculer à partir des estimations  $\tilde{a}, \tilde{b}, \tilde{c}$ ?

La réponse à cette question est oui, (à des infiniments petits près de 2ème ordre par rapport à la grandeur observée moyenne).

**Théorème 2.4** *Au deuxième ordre près, tout changement de grandeurs inconnues laisse invariant la solution des moindres carrés. Ce qui peut s'exprimer de façon plus précise ainsi :*

Etant donné un système linéaire du type :

$$\begin{aligned} {}_n A_{r \cdot r} \dot{X}_1 - K &= \dot{V}, & \text{rang } A &= r \\ B \dot{X} - M &= 0, & \text{rang } B &= p \end{aligned} \quad (2.56)$$

avec  $(r - p)$  degrés de liberté, entre l'estimateur  $\tilde{X}$  et l'estimateur  $\tilde{Y}$  qu'on obtiendrait en effectuant dans (2.56) la substitution linéaire inversible :

$${}_r \dot{X}_1 = {}_r U_{q \cdot q} \dot{Y}_1 + S \text{ avec } \begin{aligned} r &\geq q \geq r - p \\ \text{rang } U &\geq r - p \end{aligned} \quad (2.57)$$

on a la relation :

$$\boxed{{}_r \tilde{X}_1 = U {}_q \tilde{Y}_1 + S} \quad (2.58)$$

En effet dire que le système (2.57) est inversible, c'est-à-dire qu'il existe une matrice  $D$  telle que :

$${}_q \dot{Y}_1 = {}_q D_{r \cdot r} \dot{X}_1 + {}_q E_1 \iff \dot{X} = U \dot{Y} + S$$

par suite l'estimateur de  $\dot{Y}$  au sens des moindres carrés est (théorème du § 2.3.3) :

$$\tilde{Y} = D \tilde{X} + E \implies \tilde{X} = U \tilde{Y} + S$$

Il s'ensuit que quelque soit les inconnues choisies, les estimateurs au sens des moindres carrés sont bien les mêmes.

En effet, si  $\mathcal{X}$  désigne les anciennes inconnues, et  $\mathcal{Y}$  les nouvelles, on peut toujours après linéarisation se ramener à des inconnues  $\tilde{X}$  et  $\tilde{Y}$  liées linéairement, selon une relation du type (2.57) :

$$\tilde{X} = U\tilde{Y} + S$$

Ceci étant, le système (2.56) résolu, donne pour  $\tilde{X}$  l'estimateur  $\tilde{X}_1$  ; le système analogue avec des inconnues  $\tilde{Y}$  donne de celles-ci l'estimateur  $\tilde{Y}$ , et par voie de conséquence pour  $\tilde{X} = U\tilde{Y} + S$ , l'estimateur :

$$\tilde{X}_2 = U\tilde{Y} + S \quad (\text{théorème du §2.3.3})$$

mais d'après le théorème qui précède :

$$\tilde{X}_1 = U\tilde{Y} + S$$

d'où :

$$\tilde{X}_1 = \tilde{X}_2 \quad \text{ce qui entraîne } \mathcal{X}_1 = \mathcal{X}_2$$

En pratique il faut observer, qu'à l'exception du cas où on pose à priori une substitution linéaire inversible rigoureusement exacte, la substitution linéaire (2.57) résulte de la linéarisation de relations en général complexes entre anciennes et nouvelles grandeurs inconnues.

Les relations (2.57) ne sont donc vraies qu'à des infiniments petits du 2ème ordre près, par rapport à la grandeur observée moyenne; c'est pourquoi nous avons précisé dans l'énoncé du théorème, que tout changement de grandeurs inconnues laisse invariante la solution des moindres carrés, au deuxième ordre près.



### 2.3.6 *Equivalence d'un système d'équations d'observations (avec éventuellement équations de condition liant les inconnues) et d'un système d'équations de condition liant les grandeurs observées*

Il s'agit d'une conséquence directe du théorème énoncé au paragraphe précédent :

- le système d'équations d'observations (avec présence éventuelle de  $p$  équations de condition liant les inconnues) est :

$$\begin{aligned} A\hat{X} - K &= \hat{V} & \text{rang}(A) &= r \\ B\hat{X} - M &= 0, & \text{rang}(B) &= p \end{aligned} \quad (2.59)$$

Il donne des inconnues  $\hat{X}$ , l'estimateur  $\tilde{X}_1$ . - Le système des équations de condition liant les grandeurs observées  $\hat{L}$  peut s'écrire sous la forme :

$$C\hat{V} - Q = 0 \quad (2.60)$$

ou encore sous la forme :

$$\begin{aligned} \hat{Y} &= \hat{V} \\ C\hat{Y} - Q &= 0 \end{aligned} \quad (2.61)$$

Ce dernier système a une structure identique à celle du système (2.59) ; résolu par moindres carrés, il donne de  $\hat{Y}$ , l'estimateur  $\tilde{Y}$ , et par voie de conséquence en utilisant des relations appropriées l'estimateur  $\tilde{X}_2$  pour les inconnues  $\hat{X}$ .

Mais nous savons (2.2.2) que le système (2.61) se ramène au système (2.59) par la substitution linéaire inversible :

$$\hat{Y} = A\hat{X} - K \quad (2.62)$$

Il s'ensuit que (théorème § 2.3.5) :

$$\tilde{X}_1 = \tilde{X}$$

Ceci n'étant vrai, comme dans le cas étudié au § 2.3.5, qu'à des infiniments petits près du 2ème ordre par rapport aux grandeurs observées.

On peut en conclusion énoncer le théorème suivant :

**Théorème 2.5 (Théorème Fondamental)** *Tous les systèmes possibles d'équations d'observations avec équations de condition liant les inconnues, sont*

équivalents au deuxième ordre près à l'unique système des équations de condition liant les grandeurs observées (étant entendu que celui-ci est défini à un facteur matriciel régulier près).

Quelle que soit la méthode choisie, on trouve toujours au deuxième ordre près les mêmes estimateurs.

## 2.4 ESTIMATEUR AU SENS DES MOINDRES CARRÉS D'UNE FONCTION DES INCONNUES $\mathcal{Y} = f(\mathcal{X})$

On linéarise la fonction, en la développant au voisinage d'une valeur approchée  $\mathcal{X}_0$  de  $\mathcal{X}$ .

$$\begin{aligned}\mathcal{X} &= \mathcal{X} + \dot{\mathcal{X}} \\ \mathcal{Y} &= f(\mathcal{X}_0) + D\dot{\mathcal{X}} = D\dot{\mathcal{X}} + E\end{aligned}$$

Sous cette forme  $\mathcal{Y}$  apparaît comme une fonction linéaire de  $\dot{\mathcal{X}}$ ; et par suite (théorème § 2.3.3), l'estimateur de  $\mathcal{Y}$  au sens des moindres carrés est :

|  |        |
|--|--------|
| $\tilde{\mathcal{Y}} = D\tilde{\mathcal{X}} + E$ <p>avec <math>\tilde{\mathcal{X}}</math> estimateur de <math>\dot{\mathcal{X}}</math> au sens des moindres carrés</p> | (2.63) |
|--|--------|

### Application à l'exemple fondamental

Nous nous proposons d'estimer l'aire du triangle  $ABC$ , soit  $S$ , on a :

$$S = \frac{1}{2} \dot{a} \dot{c} \sin \dot{B}$$

Posant :

$$\dot{a} = a + da; \dot{c} = c_0 + dc, \left( c_0 = \frac{a \sin C}{\sin B} \right); \dot{B} = B + \dot{V}_B$$

On obtient :

$$S = \frac{1}{2} a c_0 \sin B + \frac{1}{2} c_0 \sin B . da + \frac{1}{2} a \sin B . dc + \frac{1}{2} a c_0 \cos B . \dot{V}_B \quad (2.64)$$

$da, dc$  et  $V_B$  seront calculés plus loin. Numériquement, on obtient (unité le  $mm^2$ ) :

$$S = 3.656 + 3.67985 \frac{da}{10} + 48.23977 \frac{dc}{10} + 9.50620 \frac{\pi}{2000} \tilde{V}_B \quad (2.65)$$

( $da, dc$  et  $\tilde{V}_B$  étant estimés en unités normalisées, le déci-mm et le décigrade).

## 2.5 PRATIQUE DES CALCULS

On a groupé ici un certain nombre de renseignements, de formules et de méthodes utiles dans la pratique, ainsi qu'un certain nombre d'applications numériques.

### 2.5.1 Retour sur la méthode des équations de condition liant les grandeurs observées

Le théorème fondamental de la fin du § 2.3, nous certifie que quelque soit la méthode choisie, les estimateurs des inconnues obtenus en définitive sont identiques (au 2ème ordre près).

Dans le cas d'un système d'équations de condition liant les grandeurs observées, on se trouve placé devant le système suivant, dit système des équations fondamentales :

$${}_{n-r}C_{n-n} \tilde{V}_1 - {}_{n-r}Q_1 = 0 \quad (2.66)$$

Formellement équivalent à :

$$\begin{aligned} \dot{Y} &= \dot{V} \\ C\dot{Y} - Q &= 0 \end{aligned} \quad (2.67)$$

Les estimateurs  $\tilde{Y}$  et  $\tilde{V}$  de  $\dot{Y}$  et  $\dot{V}$ , satisfont donc à :

$$\begin{aligned} \tilde{Y} &= \tilde{V} \\ C\tilde{Y} - Q &= 0 \end{aligned} \quad (2.68)$$

Il y a deux méthodes pour résoudre pratiquement un tel système, traditionnellement appelées compensation par les corrélatifs, et compensation par les éléments.

### 2.5.1.1 Compensation par les corrélatifs

On résout (2.68) par la méthode exposée en 2.3.2 (système (2.37)). Le système (2.37) s'écrit ici :

$$\begin{pmatrix} P & C^T \\ C & 0 \end{pmatrix} \cdot \begin{pmatrix} \tilde{Y} \\ \Lambda \end{pmatrix} = \begin{pmatrix} 0 \\ Q \end{pmatrix}, \quad \Lambda \text{ vecteur des corrélatifs} \quad (2.69)$$

Un tel système linéaire peut s'inverser directement, même s'il s'agit d'un système d'ordre élevé à condition dans ce dernier cas de disposer d'un programme d'inversion suffisamment élaboré (dans le cas d'observations indépendantes,  $P$  est en effet diagonale ce qui permet d'alléger notablement l'algorithme d'inversion).

On peut procéder aussi de la façon suivante : le système (2.69) équivaut à :

$$P\tilde{Y} + C^T \Lambda = 0, \quad C\tilde{Y} - Q = 0$$

d'où on déduit :

$$\tilde{Y} = -P^{-1}C^T \Lambda, \quad CP^{-1}C^T \Lambda = -Q$$

Soit enfin :

$$\boxed{\begin{aligned} \Lambda &= -(CP^{-1}C^T)^{-1}Q \\ \tilde{Y} &= P^{-1}C^T(CP^{-1}C^T)^{-1}Q = -P^{-1}C^T \Lambda \end{aligned}} \quad (2.70)$$

### 2.5.2 Compensation par les éléments

Etant donné le système d'équations de condition écrit sous la forme :

$$\begin{aligned} \tilde{Y} &= \tilde{V} \\ {}_{n-r}C_{nn}\tilde{Y}_1 - {}_{n-r}Q_1 &= 0 \end{aligned} \quad (2.71)$$

on peut aussi, à partir de la deuxième équation, exprimer  $(n-r)$  des composantes de  $Y$ , que nous supposons être les  $(n-r)$  dernières en fonction des  $r$  premières :

$$C\tilde{Y} - Q = 0 \implies {}_{n-r}\tilde{Y}_1^{(n-r)} = C_{1,r}\tilde{Y}_1^{(r)} + Q_1 \quad (2.72)$$

tenant compte de ceci, la première équation de (2.71) se met sous la forme :

$$A_{,r}\tilde{Y}_1^{(r)} - K = \tilde{V} \quad (2.73)$$

Les  $r$  inconnues  $\tilde{Y}^{(r)}$  s'appellent traditionnellement éléments.

Après résolution par moindres carrés, on obtient :

$$\tilde{Y}^{(r)} = (A^T P A)^{-1} A^T P K \quad (2.74)$$

d'où  $\tilde{Y}^{(n-r)}$  par (2.72). Puis  $\tilde{V} = \tilde{Y}$ .

Cette méthode qui nécessite avant calcul des manipulations algébriques (2.72-2.73), semble assez rarement utilisée.

### 2.5.3 Formulaire

**Méthode des équations d'observations** ( $r$  grandeurs indépendantes à déterminer)

On exprime les  $n$  grandeurs observées  ${}_n L_1$  en fonction des  $r$  grandeurs inconnues  ${}_r \mathcal{X}_1$  indépendantes :

$$L = \varphi(\mathcal{X}) = \varphi(\mathcal{X}_0 + \dot{X})$$

Linéarisation :

$${}_n A_{rr} \dot{X}_1 - {}_n K_1 = {}_n \dot{V}_1$$

$$\text{rang} A = r$$

$$\tilde{V}^T P \tilde{V} \text{ minimal}$$

⇒

$$\text{Système normal : } A^T P A \tilde{X} = A^T P K$$

$$\text{Matrice normale : } {}_r N_r = A^T P A$$

$$\text{Solution : } \tilde{X} = N^{-1} A^T P K$$

$$\text{Résidus : } \tilde{V} = A \tilde{X} - K$$

$$\text{Observations compensées : } \tilde{L} = L + \tilde{V}$$

$$\text{Vérification : } A^T P \tilde{V} = 0$$

$$\text{Estimation }^a \text{ de } \sigma_0^2 : s_0^2 = \frac{\tilde{V}^T P \tilde{V}}{n-r}$$

(variance unitaire)

Matrice de variance <sup>b</sup> de  $\tilde{X}$  :

$$\Gamma_{\tilde{X}} = N^{-1} \sigma_0^2 \tilde{\Gamma}_{\tilde{X}} = N^{-1} s_0^2$$

a. Ces expressions seront établies dans le chapitre 4.

b. Idem.

**Méthode des équations de condition liant les grandeurs observées** $(r$  grandeurs inconnues, liées par  $p$  relations)

On exprime les  $n$  grandeurs observées  ${}_n\dot{L}_1$  en fonction des  $r$  grandeurs inconnues  $\tilde{X}$  et on écrit les  $p$  équations de condition liant les inconnues :

$$\begin{aligned} {}_n\dot{L}_1 &= \varphi(\tilde{X}) = \varphi(\tilde{X}_0 + \dot{X}) \\ \psi_i(\tilde{X}_0 + \dot{X}) &= 0; (i = 1, p) \end{aligned}$$

Linéarisation :

$$\begin{aligned} {}_nA_{rr}\dot{X}_1 - {}_nK_1 &= {}_n\dot{V}_1 \\ \text{rang}A &= r \\ {}_pB_{rr}\dot{X}_1 - {}_pM_1 &= 0 \\ \text{rang}B &= p \end{aligned}$$

 $\Rightarrow$ 

$$\tilde{V}^T P \tilde{V} \text{ minimal}$$

$$\text{Système normal : } \begin{pmatrix} A^T P A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \tilde{X} \\ \Lambda \end{pmatrix} = \begin{pmatrix} A^T P K \\ M \end{pmatrix}$$

Vecteur des corrélatifs :  ${}_p\Lambda_1$ 

$$\text{Vérification : } A^T P \tilde{V} + B^T \Lambda = 0$$

$$\text{Estimation }^a \text{ de } \sigma_0^2 \text{ (variance unitaire) : } s_0^2 = \frac{\tilde{V}^T P \tilde{V}}{n - (r - p)}$$

$$\tilde{V}^T P \tilde{V} = K^T P K - \tilde{X}^T N \tilde{X} - 2B^T \Lambda$$

$$\text{Matrice variance }^b \text{ de } X : \Gamma_{\tilde{X}} = (\delta_{ij}) \cdot \begin{pmatrix} A^T P A & B^T \\ B & 0 \end{pmatrix}^{-1} \cdot \sigma_0^2$$

$$\delta_{ij} = \begin{cases} = 1 & \text{si } i = j \leq r \\ = 0 & \text{pour les autres cas} \end{cases}$$

*a.* Expressions établies dans le chapitre 4.

*b.* Idem.

**Méthode des équations de condition liant les grandeurs observées**

Si  $r$  est le nombre de degrés de liberté de la figure (déterminée par  $r$  grandeurs indépendantes  ${}_r\mathcal{X}_1$ ) on écrit que les  $n$  grandeurs observées  ${}_n\dot{L}_1$  sont liées par  $(n - r)$  rela-

tions indépendantes :

$$\Psi_i(L) = \Psi_i(L + \dot{V}) = 0, i = 1, n - r$$

Linéarisation :

$${}_{n-r}C_{nn}\dot{V}_1 - {}_{n-r}Q_1 = 0$$

$$\text{rang } C = n - r$$

$$\tilde{V}^T P \tilde{V} \text{ minimal}$$

$\Rightarrow$

$$\text{Système normal : } \begin{pmatrix} P & C^T \\ C & 0 \end{pmatrix} \begin{pmatrix} \tilde{V} \\ \Lambda \end{pmatrix} = \begin{pmatrix} 0 \\ Q \end{pmatrix}$$

$$\text{Vecteur des corrélatifs : } \Lambda = -(CP^{-1}C^T)^{-1}Q$$

$$\text{Solution : } \tilde{V} = -P^{-1}C^T\Lambda$$

$$\text{Vérification : } P\tilde{V} + C^T\Lambda = 0$$

$$\text{Observations compensées : } \tilde{L} = L + \tilde{V} \Rightarrow \tilde{\mathcal{X}}$$

Estimation<sup>a</sup> de  $\sigma_0^2$  (variance unitaire) :

$$s_0^2 = \frac{\tilde{V}^T P \tilde{V}}{n - r}$$

$$\text{Matrice variance}^b \text{ des résidus } V : \Gamma_{\tilde{V}} = (\delta_{ij}) \cdot \begin{pmatrix} P & C^T \\ C & 0 \end{pmatrix}^{-1} \cdot \sigma_0^2$$

$$\delta_{ij} = \begin{cases} = 1 & \text{si } i = j \leq r \\ = 0 & \text{pour les autres cas} \end{cases}$$

a. Expressions établies dans le chapitre 4.

b. Idem.

**Estimation d'une fonction linéaire des inconnues**

$$\mathcal{Y} = f(\mathcal{X}) = f(\mathcal{X}_0 + \dot{X})$$

Linéarisation :

$$\dot{\mathcal{Y}} = D\dot{X} + E$$

 $\implies$ 

$$\text{Solution : } \tilde{\mathcal{Y}} = D\tilde{X} + E$$

Matrice de variance de  $\tilde{Y}$  :

$$\Gamma_{\tilde{Y}} = D\Gamma_{\tilde{X}}D^T$$

**2.5.4 Rappel sur les méthodes d'inversion d'un système linéaire régulier****2.5.4.1 Rappels des principes et algorithmes relatifs à deux méthodes directes**

Les méthodes directes sont par définition des méthodes qui donnent en principe la solution exacte d'un système linéaire moyennant un nombre fini d'opérations élémentaires (+, -, x, /,  $\sqrt{\quad}$ ); leur coût dépend donc du nombre de ces opérations élémentaires.

La solution calculée diffère de la solution exacte par suite de l'accumulation des erreurs d'arrondi : on dit qu'une méthode est stable si elle est peu sensible à l'accumulation des erreurs d'arrondi.

Les méthodes directes permettent de résoudre des systèmes pleins d'ordre  $\leq 350$  à condition de disposer d'un ordinateur puissant; mais dans le cas de systèmes très creux, on peut résoudre des systèmes d'ordre beaucoup plus élevés (jusqu'à 15000 inconnues et plus); tel est le cas lorsqu'on "compense" des réseaux, où chaque sommet de maille n'est lié par mesures qu'à un très petit nombre de sommets voisins : les systèmes sont alors fortement diagonalisés.

Nous ne nous intéresserons dans ce qui suit qu'aux systèmes carrés symétriques réguliers :

$${}_nN_{nn}X_1 - {}_nW_1 = 0$$



### 2.5.4.2 Principe de la résolution par triangulation d'un système carré symétrique régulier ${}_nN_{nn}X_1 - {}_nW_1 = 0$

Les méthodes de triangularisation sont infiniment plus économiques et rapides que la méthode qui consisterait à utiliser les formules de Cramer (environ  $n^2n!$  soit  $10^8$  opérations élémentaires pour  $n = 10$  dans ce cas, contre un nombre de l'ordre de  $n^3$  pour les méthodes de triangularisation. . .).

Les deux méthodes les plus utilisées sont celles de Gauss-Doolittle et celle de Cholevsky; elles comportent chacune deux phases analogues :

- une triangularisation du système, appelée aussi descente :

$$NX - W = 0 \implies RX - W' = 0$$

avec : ( $R$  est dite triangulaire supérieure droite).

$$R = \begin{vmatrix} \times & & \\ & \times & \\ & & \times \end{vmatrix}$$

- la résolution du système triangularisé, appelée aussi remontée :

$$RX - W' = 0 \implies X = R^{-1}W'$$

L'algorithme appliqué à  $W$ , nous donne ainsi  $X$ ; si on désire de surcroît obtenir la matrice inverse  $N^{-1}$  (calcul de matrice de variance), il faut résoudre  $n$  systèmes linéaires du type :

$${}_nN_n X_i = V_i \quad \text{avec} \quad {}_nV_{1i} = (\delta_{ij})$$

On obtiendra ainsi successivement tous les vecteurs colonnes de  $N^{-1}$ .

### 2.5.4.3 Résolution d'un système triangularisé $RX - W' = 0$ : algorithme de la remontée

Cet algorithme est immédiat à établir; le système  $RX - W' = 0$  s'écrit :

$$\begin{aligned} r_{11}x_1 + r_{12}x_2 + \cdots + r_{1,n-1}x_{n-1} + r_{1n}x_n &= w'_1 \\ r_{22}x_2 + \cdots + r_{2,n-1}x_{n-1} + r_{2n}x_n &= w'_2 \end{aligned}$$

$$\begin{aligned}
 & \dots\dots\dots \\
 & r_{ii}x_i + \dots\dots\dots + r_{i,n-1}x_{n-1} + r_{in}x_n = w'_i \\
 & \dots\dots\dots \\
 & r_{n-1,n-1}x_{n-1} + r_{n-1,n}x_n = w'_{n-1} \\
 & r_{nn}x_n = w'_n
 \end{aligned}$$

D'où on déduit :

$$\begin{aligned}
 & x_n = w'_n / r_{nn} \\
 & x_{n-1} = (w'_{n-1} - r_{n-1,n}x_n) / r_{n-1,n-1} \\
 & \vdots \\
 & x_i = (w'_i - r_{in}x_n - r_{i,n-1}x_{n-1} \dots - r_{i,i-1}x_{i+1}) / r_{ii} \\
 & \vdots
 \end{aligned} \tag{2.75}$$

On remarquera que les pivots sont différents de 0 (sans quoi le système  $RX - W' = 0$  ne serait pas résoluble).

**2.5.4.4 Descente par l'algorithme de Gauss-Doolittle**

Etant donné :  $NX - W = 0$ , on pose :

|   |        |
|---|--------|
| $N = \mathcal{R}.R \quad \text{avec : } \begin{matrix} \mathcal{R} & \text{matrice triangulaire inférieure} \\ R & \text{matrice triangulaire supérieure} \end{matrix}$ | (2.76) |
|---|--------|

On obtient alors successivement :  $RX = \mathcal{R}^{-1}W = W'$ .

C'est la descente. Puis :  $X = R^{-1}W'$ .

C'est la remontée.

Les différentes étapes du calcul sont figurées dans le schéma ci-après; les parties non hachurées des tableaux sont creuses; de plus on tient compte dans la suite des calculs de la symétrie des matrices  $N, N^{-1}$ , et du fait que  $R$  est triangulaire.

On notera, remarque intéressante pour la programmation, que si on ne désire pas calculer la matrice inverse  $N^{-1}$ , tous les calculs peuvent être effec-

tués sur un seul tableau triangulaire de  $n(n + 1)/2$  mémoires avec en plus une colonne de  $n$  mémoires pour le terme constant.

Les contrôles, en cas de calcul manuel, se font en appliquant l’algorithme au vecteur somme  $\Sigma$ .

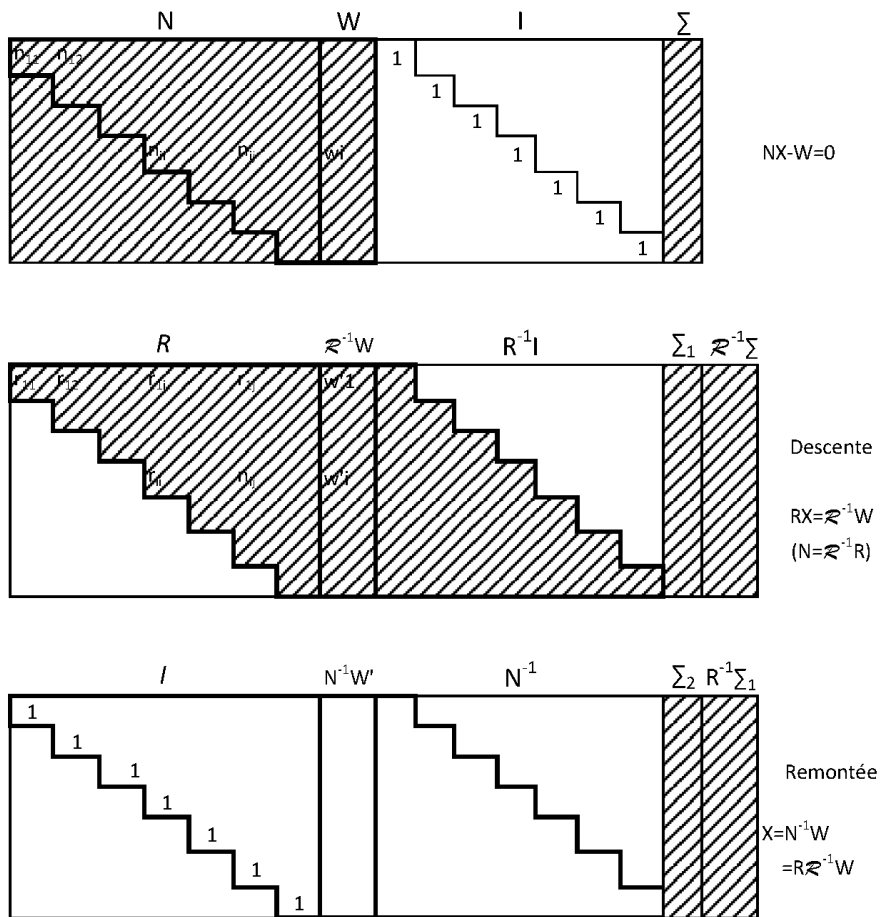


Fig. 2.2 Etapes du calcul de la descente par l’algorithme de Gauss-Doolittle

**Algorithme de la descente - GAUSS-DOOLITTLE<sup>2</sup>**1er pivot:  $r_{11} = n_{11}$ 1ère ligne de  $R$ :  $r_{1j} = n_{1j}$ 1ère composante de  $W' = \mathcal{R}^{-1}W$ :  $w'_1 = w_1$ 

$$\begin{aligned}
 i\text{ème pivot: } r_{ii} &= n_{ii} - \sum_{k=1}^{i-1} \frac{r_{ki}^2}{r_{kk}} \\
 i\text{ème ligne de } R: r_{ij} &= n_{ij} - \sum_{k=1}^{i-1} \frac{r_{kj}r_{ki}}{r_{kk}} \\
 i\text{ème composante de } W': w'_i &= w_i - \sum_{k=1}^{i-1} \frac{w'_k r_{ki}}{r_{kk}}
 \end{aligned}
 \tag{2.77}$$

**Example :**

$$\begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} X = \begin{vmatrix} 3 \\ 1 \end{vmatrix}$$

|   |    |      |      |      |    |    |
|---|----|------|------|------|----|----|
| 1 | 2  | 3    | 1    |      | 7  |    |
|   | 1  | 1    |      | 1    | 5  |    |
| 1 | 2  | 3    | 1    |      | 7  | 7  |
|   | -3 | -5   | -2   | 1    | -9 | -9 |
| 1 |    | -1/3 | -1/3 |      | 1  | 1  |
|   | 1  | 5/3  | 2/3  | -1/3 | 3  | 3  |

**Contrôle**

$$X = (-1/3, 5/3)^T$$

$$\begin{vmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{vmatrix} = N^{-1}$$

**Propriétés de l'algorithme de Gauss-Doolittle**

- Il y a environ  $2n^3/3$  opérations élémentaires.
- On ne peut garantir la stabilité de la méthode sans stratégie convenable sur le choix des pivots successifs.

On montre en effet qu'on risque, si on opère sans précautions de tomber sur des pivots très petits, et par suite de rendre les calculs très imprécis, car les conséquences des erreurs d'arrondi dont l'ordre de grandeur est constant

2. On montre sans peine que cet algorithme reproduit la méthode des éliminations successives.

sont beaucoup plus graves sur de petits nombres que sur de grands nombres, soit, par exemple :  $u = \frac{N}{D}$  avec  $N$  et  $D$  petits.

Si  $\varepsilon_1$  et  $\varepsilon_2$  sont les erreurs d'arrondi sur  $N$  et  $D$ , on a :

$$\left| \frac{du}{u} \right| = \left| \frac{\varepsilon_1}{N} - \frac{\varepsilon_2}{D} \right|$$

alors que si on prend :

$$n = \frac{kN}{kD} \quad \text{avec} \quad D \ll kD$$

$$\text{on a : } \left| \frac{du}{u} \right| = \left| \frac{\varepsilon'_1}{kN} - \frac{\varepsilon'_2}{kD} \right| \ll \left| \frac{\varepsilon_1}{N} - \frac{\varepsilon_2}{D} \right|$$

Les deux stratégies de pivots les plus utilisées pour l'algorithme de Gauss sont :

- la stratégie par pivot partiel :

On élimine lors de la descente les inconnues dans l'ordre naturel en permutant les lignes qui restent à descendre de façon à faire apparaître le pivot maximal.

- la stratégie par pivot total :

On permute dans la matrice résiduelle lignes et colonnes de façon à opérer avec le pivot maximal : cela revient à déterminer dans la matrice résiduelle l'élément de plus grand module et à effectuer sur les inconnues et les termes constants résiduels et les permutations adéquates.

Ces deux stratégies garantissent la stabilité de la méthode, c'est-à-dire minimisent l'influence de l'accumulation des erreurs arrondi.

#### 2.5.4.5 Descente par l'algorithme de Cholewsky

Etant donné :  $NX - W = 0$ , on pose :

$$N = R^T R \quad \text{avec : } R \text{ matrice triangulaire supérieure}$$

On obtient alors (descente) :  $RX - (R^T)^{-1}W = 0$

puis (remontée) :  $X = R^{-1}(R^T)^{-1}W$ .

D'où le schéma suivant (mêmes conventions que pour l'algorithme de Gauss-Doolittle).

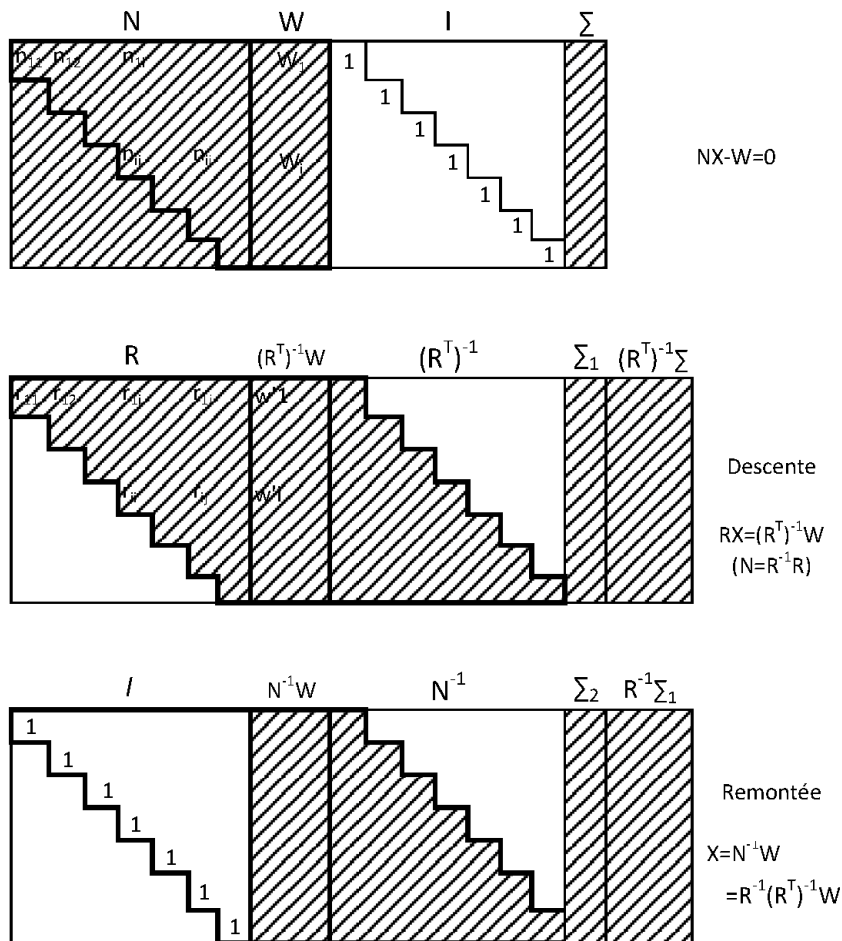


Fig. 2.3 Etapes du calcul de la descente par l'algorithme de Cholesky

**Algorithme de la descente de Cholewsky**

$$\text{1er pivot : } r_{11} = \sqrt{n_{11}}$$

$$\text{1ère ligne de } R : r_{1j} = \frac{n_{1j}}{r_{11}}$$

$$\text{1ère composante de } W' = (R^T)^{-1}W : w'_1 = \frac{w_1}{r_{11}}$$

$$\begin{aligned} i\text{ème pivot : } r_{ii} &= \sqrt{n_{ii} - \sum_{k=1}^{i-1} r_{ki}^2} \\ i\text{ème ligne de } R : r_{ij} &= \left( n_{ij} - \sum_{k=1}^{i-1} r_{kj}r_{ki} \right) / r_{kk} \\ i\text{ème composante de } W' : w'_i &= \left( w_i - \sum_{k=1}^{i-1} w'_k r_{ki} \right) / r_{kk} \end{aligned} \quad (2.78)$$

N.B. Si on a à extraire la racine carrée d'un nombre négatif  $A$  on prendra  $\sqrt{A} = i\sqrt{-A}$  avec  $i$  le nombre complexe  $\sqrt{-1}$ .

**Exemple :**

|   |            |              |              |              |              |              |
|---|------------|--------------|--------------|--------------|--------------|--------------|
| 1 | 2          | 3            | 1            |              | 7            |              |
|   | 1          | 1            |              | 1            | 5            |              |
| 1 | 2          | 3            | 1            |              | 7            | 7            |
|   | $i3^{1/2}$ | $i5/3^{1/2}$ | $i2/3^{1/2}$ | $-i/3^{1/2}$ | $i9/3^{1/2}$ | $i9/3^{1/2}$ |
| 1 |            | -1/3         | -1/3         |              | 1            | 1            |
|   | 1          | 5/3          | 2/3          | -1/3         | 3            | 3            |

Contrôle

$$X = (-1/3, 5/3)^T$$

$$N^{-1} = \begin{vmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{vmatrix}$$

### **Propriétés de l'algorithme de Cholewsky**

- Il y a environ  $n^3/3$  opérations élémentaires, pour la résolution d'un système de type  $NX - W = 0$ ; le gain est donc très sensible par rapport à l'algorithme de Gauss-Doolittle.

- A condition que la matrice soit symétrique définie positive, l'algorithme est numériquement stable sans qu'il soit nécessaire d'effectuer des permutations de lignes ou d'inconnues, avantage considérable également sur l'algorithme de Doolittle; noter qu'on ne peut plus garantir la stabilité, si la matrice n'est pas symétrique, définie positive.

- Si  $N$  est  $(2p+1)$  diagonale ( $n_{ij} = 0$ ) pour  $|i-j| \geq p+1$ , il en est de même pour les matrices  $R$  et  $N^{-1}$ ; autrement dit l'algorithme de Cholewsky préserve la structure bande des matrices, d'où condensation possible de mémoires. (Ce qui n'est vrai pour l'algorithme de Gauss que si on ne permute ni lignes, ni colonnes; mais alors la méthode n'est plus stable).

Ainsi le seul inconvénient de la méthode est d'obliger à l'extraction de racines carrées; mais ce n'en est plus un pour le calcul sur ordinateur.

#### **2.5.4.6 Application numérique**

Sont traités ci-dessous les trois cas suivants :

- l'exemple fondamental par la méthode de l'équation d'observation (A2; matrice poids A3; système des équations d'observations (2.14)).
- l'exemple fondamental par la méthode des équations de condition (2.29).
- l'exemple fondamental avec une équation liant les grandeurs inconnues (angle  $B = \pi/2$ ) (2.18).

#### **Exemple fondamental par la méthode des relations d'observations**

Le système (2.14) est sous la forme :





$$\tilde{V} = \begin{pmatrix} \tilde{V}_a = 0.62971 \\ \tilde{V}_b = -0.90962 \\ \tilde{V}_A = 0.41701 \\ \tilde{V}_B = 0.53027 \\ \tilde{V}_C = 0.53105 \end{pmatrix}, \quad \sqrt{P}\tilde{V} = \begin{pmatrix} 0.33123 \\ -0.36385 \\ 0.51501 \\ 0.65488 \\ 0.65585 \end{pmatrix}$$

$$A^T P \tilde{V} = \begin{pmatrix} 0.00002 \\ -0.00002 \\ 0.00001 \end{pmatrix} \quad \text{Contrôle normalisation et résolution}$$

$$\tilde{V}^T P \tilde{V} = 1.36634$$

$$\text{Côtés compensés : } \tilde{\mathcal{X}} = \begin{pmatrix} \tilde{a} = 965.42971 \\ \tilde{b} = 1154.09038 \\ \tilde{c} = 634.54782 \end{pmatrix}$$

$$\text{Angles compensés : } \begin{pmatrix} \tilde{A} = 630.83701 \\ \tilde{B} = 998.55027 \\ \tilde{C} = 370.61105 \end{pmatrix} \implies \tilde{A} + \tilde{B} + \tilde{C} = 1999.99833 \text{ dcgr}$$

$$s_0^2 = \frac{\tilde{V}^T P \tilde{V}}{n-r} = \frac{1.36634}{5-3} = 0.68317 \implies s_0 = 0.82654$$

### Exemple fondamental par la méthode des équations de condition

En exprimant que les  $n$  grandeurs observées sont liées par  $n-r$  relations indépendantes, on a obtenu (2.2.2) le système :

$$C\dot{V} - Q = 0$$

auquel on peut substituer :

$$\begin{aligned} \dot{Y} &= \dot{V} \\ C\dot{Y} - Q &= 0 \end{aligned}$$

Nous résoudrons ce système par la méthode des corrélatifs (2.5.1.1);  $\tilde{V}$  et le vecteur des corrélatifs  $\Lambda$  sont solutions de :

$$\begin{pmatrix} P & C^T \\ C & 0 \end{pmatrix} \begin{pmatrix} \tilde{V} \\ \Lambda \end{pmatrix} = \begin{pmatrix} 0 \\ Q \end{pmatrix} \quad (2.79)$$

qui s'écrit aussi :

$$\mathcal{N} \cdot \begin{pmatrix} \tilde{V} \\ \Lambda \end{pmatrix} = \begin{pmatrix} 0 \\ Q \end{pmatrix}$$

On peut alors soit utiliser (2.70), soit résoudre directement, c'est-à-dire inverser le système (2.79); on voit aisément que la masse des calculs est du

même ordre dans les deux cas.

Il est en outre intéressant de calculer  $\mathcal{N}^{-1}$  : nous verrons en effet que la sous-matrice supérieure gauche d'ordre  $n$  de  $\mathcal{N}^{-1}$  est la matrice de variance  $\Gamma_{\tilde{V}}$  de  $\tilde{V}$  au facteur  $\sigma_0^2$  près.

On constate que les résidus  $V$  obtenus ne sont identiques à ceux calculés par la méthode des équations d'observations qu'à  $10^{-3}$  près (alors que la vérification des calculs se fait à  $10^{-5}$  près) : il faut y voir l'influence des termes négligés dans la linéarisation.

Les calculs figurent sur la page suivante (**Fig. 2.4**).

On obtient :

$$\begin{pmatrix} \tilde{V} \\ \Lambda \end{pmatrix} = \begin{pmatrix} \tilde{V}_a = 0.63090 \\ \tilde{V}_b = -0.91145 \\ \tilde{V}_A = 0.41716 \\ \tilde{V}_B = 0.53170 \\ \tilde{V}_C = 0.53116 \\ \Lambda_1 = -0.80948 \\ \Lambda_2 = -0.20900 \end{pmatrix}, \quad \sqrt{P}\tilde{V} = \begin{pmatrix} 0.33205 \\ -0.36458 \\ 0.51498 \\ 0.65138 \\ 0.65572 \end{pmatrix}$$

$$P\tilde{V} + C^T \Lambda = \begin{pmatrix} -0.00000 \\ -0.00000 \\ -0.00000 \\ 0.00001 \\ 0.00001 \end{pmatrix} \quad \text{Contrôle normalisation et résolution}$$

$\tilde{V}^T P \tilde{V} = 1.36918 \quad s_0 = 0.82740, \Gamma_{\tilde{V}} / \sigma_0^2 =$  sous-matrice (5,5) supérieure gauche de  $\mathcal{N}^{-1}$

$$\text{Angles compensés : } \begin{pmatrix} \tilde{A} = 630.83316 \\ \tilde{B} = 998.55170 \\ \tilde{C} = 370.61116 \end{pmatrix} \implies \tilde{A} + \tilde{B} + \tilde{C} = 2000.00021 \text{ dcgr}$$

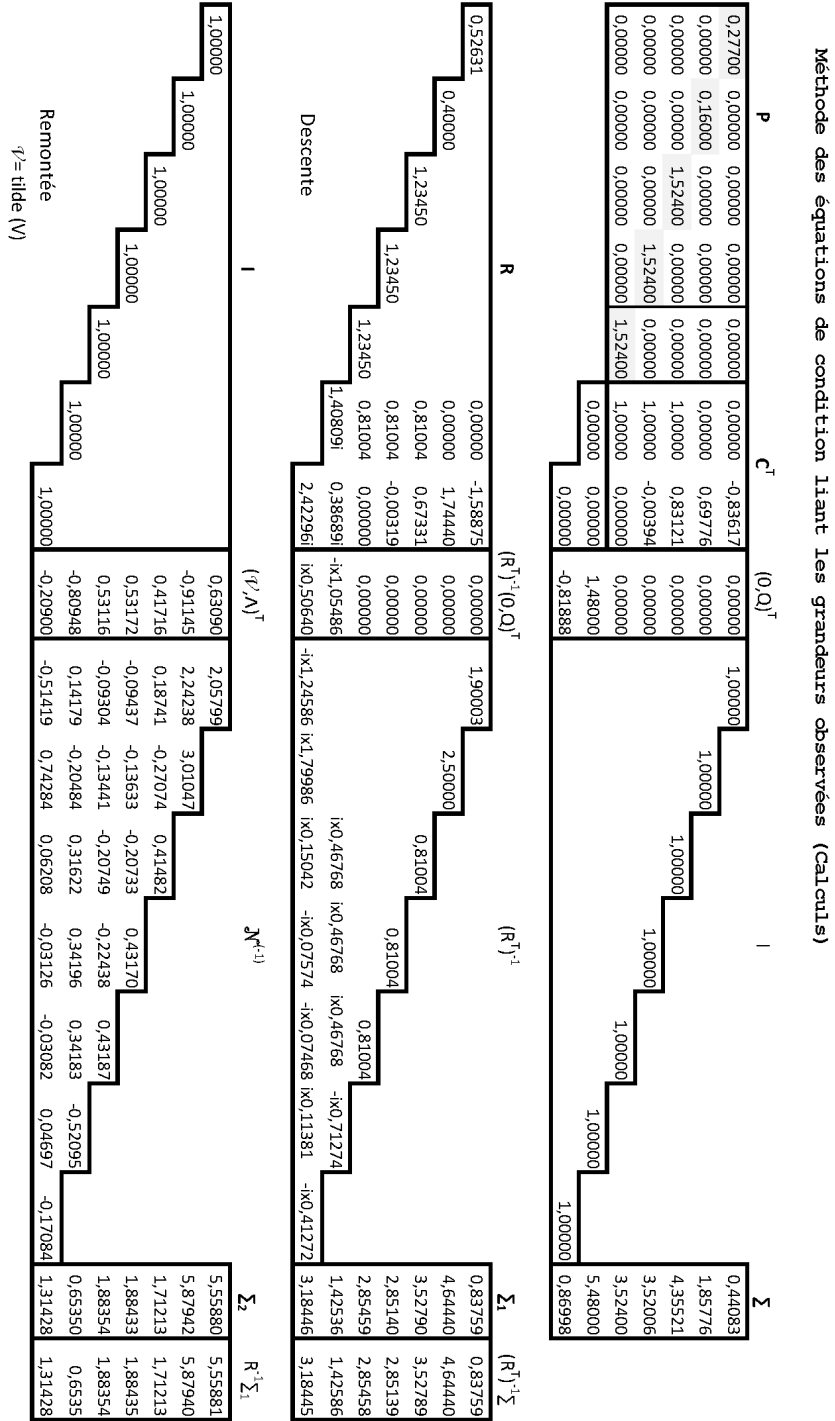


Fig. 2.4 Méthode des équations de condition liant les grandeurs observées

**Exemple fondamental avec la méthode des équations d'observations  
mais avec une équation de condition liant les grandeurs inconnues**

En imposant à l'angle  $B$  d'être droit, on a été amené à joindre au système des équations d'observations :

$$A\tilde{X} - K = \tilde{V}$$

une équation de condition de la forme :

$$B\tilde{X} - M = 0$$

L'estimateur au sens des moindres carrés  $\tilde{X}$  de  $\tilde{X}$  est alors solution du système :

$$\begin{pmatrix} A^T P A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \tilde{X} \\ \Lambda \end{pmatrix} = \begin{pmatrix} A^T P K \\ M \end{pmatrix}$$

qu'on peut écrire :

$$\mathcal{N}^{-1} \begin{pmatrix} \tilde{X} \\ \Lambda \end{pmatrix} = \begin{pmatrix} A^T P K \\ M \end{pmatrix} \quad (2.80)$$

Là encore, on a résolu directement le système (2.80) sans élimination intermédiaire, et calculé  $\mathcal{N}^{-1}$ ; nous verrons que la sous-matrice  $3 \times 3$  supérieure gauche de  $\mathcal{N}^{-1}$  est la matrice variance de  $\tilde{X}$  au facteur  $\sigma_0^2$  près.

On trouve un angle  $B$  égal à  $\pi/4$  à 4 secondes centé près; il est d'ailleurs probable que ce résidu de 4 secondes ne provient pas des erreurs de calcul (précis au moins à  $10^{-4}$  près) mais des termes négligés dans la linéarisation.

On obtient :

$$\tilde{V} = \begin{pmatrix} \tilde{V}_a = -0.32776 \\ \tilde{V}_b = -0.44061 \\ \tilde{V}_A = 0.28095 \\ \tilde{V}_B = 1.97961 \\ \tilde{V}_C = -0.21886 \end{pmatrix}, \quad \sqrt{P}\tilde{V} = \begin{pmatrix} 0.17240 \\ -0.17624 \\ -0.34697 \\ 2.44482 \\ -0.27029 \end{pmatrix}$$

$$A^T P \tilde{V} + B^T \Lambda = \begin{pmatrix} 0.00001 \\ -0.00001 \\ 0.00001 \end{pmatrix} \quad \text{Contrôle normalisation et résolution}$$

$$\tilde{V}^T P \tilde{V} = 6.18972 \quad s_0^2 = \frac{1}{5 - (3 - 1)} \tilde{V}^T P \tilde{V} = 2.06224$$

$$\Gamma_{\tilde{X}} / \sigma_0^2 = \text{sous-matrice (3,3) supérieure gauche de } \mathcal{N}^{-1}$$

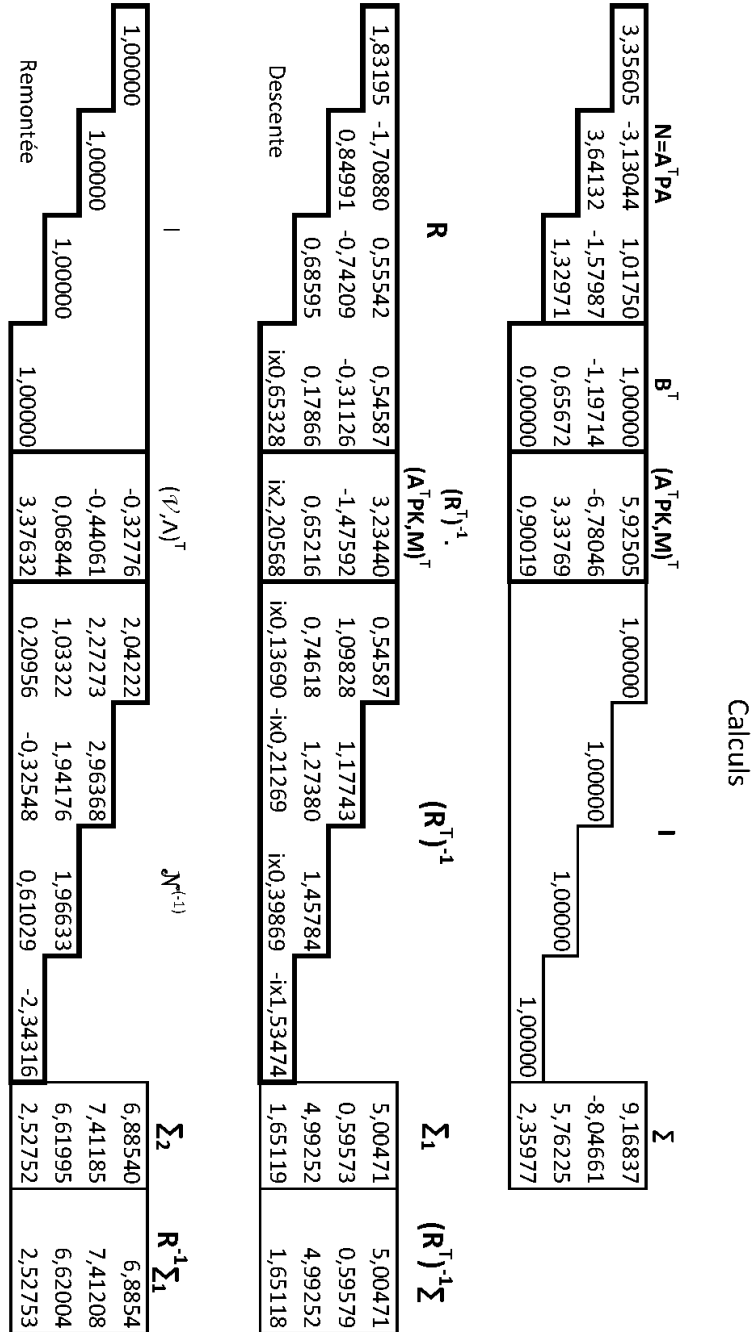


Fig. 2.5 Equations d'observations + équations de condition

$$\text{Côtés compensés : } \tilde{\mathcal{X}} = \begin{pmatrix} \tilde{a} = 965.12776 \\ \tilde{b} = 1154.55939 \\ \tilde{c} = 633.66901 \end{pmatrix}$$

$$\text{Angles compensés : } \begin{pmatrix} \tilde{A} = 630.13905 \\ \tilde{B} = 999.99961 \\ \tilde{C} = 369.86114 \end{pmatrix} \implies \tilde{A} + \tilde{B} + \tilde{C} = 1999.99980 \text{ } dcgr$$





# Chapitre 3

## Compléments Sur La Théorie Elémentaire Des Moindres Carrés

### Sommaire

---

|       |   |    |
|-------|---|----|
| 3.1   | FORME CANONIQUE D'UN SYSTÈME D'ÉQUATIONS<br>D'OBSERVATIONS .....  | 62 |
| 3.2   | ESTIMATION $\tilde{X}$ D'UNE GRANDEUR À $r$ DIMENSIONS $\tilde{X}$ À<br>PARTIR D'UN ÉCHANTILLON DE $N$ ESTIMATIONS $X_1, X_2, \dots, X_N$<br>DE POIDS $P_1, P_2, \dots, P_N$ ( $P_i = \text{matrice}$ ) ..... | 63 |
| 3.3   | CONTRIBUTIONS DE CHAQUE ÉQUATION D'OBSERVATIONS,<br>OU D'UNE PARTIE DU SYSTÈME DES ÉQUATIONS<br>D'OBSERVATIONS, À LA MATRICE NORMALE DANS LE CAS OÙ<br>LES OBSERVATIONS SONT INDÉPENDANTES .....              | 64 |
| 3.4   | PROBLÈME DES OBSERVATIONS RÉPÉTÉES- CHOIX DES<br>POIDS POUR UN SYSTÈME DE MESURES ENTIÈREMENT<br>DÉTERMINÉ- CHOIX DES POIDS POUR DES OBSERVATIONS<br>SUPPLÉMENTAIRES D'UNE GRANDEUR DÉJÀ OBSERVÉE .....       | 65 |
| 3.4.1 | Groupement des observations relatives à une même<br>grandeur observée .....   | 65 |
| 3.4.2 | Etude d'un système de mesures entièrement<br>déterminé .....  | 68 |
| 3.4.3 | Choix des poids d'observations supplémentaires<br>relatives à une grandeur observée .....   | 69 |

|            |  |           |
|------------|--|-----------|
| <b>3.5</b> | <b>ADJONCTION D'OBSERVATIONS OU D'INCONNUES<br/>SUPPLÉMENTAIRES (cas de grands réseaux) - MÉTHODE DU<br/>BORDAGE .....</b> | <b>71</b> |
|------------|--|-----------|

### 3.1 FORME CANONIQUE D'UN SYSTÈME D'ÉQUATIONS D'OBSERVATIONS

Le système le plus général que nous avons considéré est le suivant :

$$\begin{aligned} {}_n A_{rr} \dot{X}_1 - {}_n K_1 &= {}_n \dot{V}_1 \quad \text{rang } A = r \\ {}_p B_{rr} \dot{X}_1 - {}_p M_1 &= 0 \quad \text{rang } B = p \end{aligned} \quad (3.1)$$

Il est constitué d'un système de  $n$  équations d'observations et d'un système de  $p$  équations de condition liant les grandeurs inconnues.

La matrice de poids est  $P$ ; elle n'est pas forcément diagonale (ainsi les mesures indirectes auxquelles on assimile les résultats  $\tilde{X}$  d'une compensation amont, ont pour matrice de poids une matrice non diagonale en général).

La solution  $\tilde{X}$  des moindres carrés est obtenue en minimisant la forme quadratique fondamentale  $\tilde{V}^T P \tilde{V}$ ; il est immédiat de vérifier que le système (3.1) joint à la matrice de poids  $P$  équivaut au système suivant, joint à une matrice de poids égale à la matrice unité :

$$\begin{array}{l} \sqrt{P} A \dot{X} - \sqrt{P} K = \dot{V}_1 \\ B \dot{X} - M = 0 \quad \text{nouvelle matrice de poids : } I \end{array} \quad (3.2)$$

Nous dirons alors que le système initial des équations d'observations a été mis sous forme canonique. Il est clair que :

$$\dot{V} \in \mathcal{N}(0, P^{-1/2} \sigma_0) \implies \dot{V}_1 \in \mathcal{N}(0, \sigma_0)$$

autrement dit les observations du système (3.2) sont indépendantes, de même écart-types et si les observations initiales sont normales, normales.

**Définition 3.1** Nous appellerons canonique un système d'équations d'observations où les observations sont indépendantes et de même écart-type; il est

*toujours possible de passer d'une forme non canonique à une forme canonique : il suffit de multiplier les deux membres par  $\sqrt{P}$  (racine carrée de la matrice des poids).*

*En particulier lorsque les mesures sont directes donc indépendantes, et hétérogènes, écrire le système sous forme canonique revient à diviser chaque équation par l'écart-type de l'observation, c'est-à-dire à prendre comme unité de la mesure correspondante son écart-type.*

### 3.2 ESTIMATION $\tilde{X}$ D'UNE GRANDEUR À $r$ DIMENSIONS $\tilde{X}$ À PARTIR D'UN ÉCHANTILLON DE $N$ ESTIMATIONS $X_1, X_2, \dots, X_N$ DE POIDS $P_1, P_2, \dots, P_N$ ( $P_i = \text{matrice}$ )

Ce problème généralise celui de l'estimation d'une grandeur à une dimension à partir de  $n$  mesures pondérées,  $X_1, X_2, \dots, X_n$  de poids  $p_1, p_2, \dots, p_n$  (scalaires).

On avait trouvé alors que le meilleur estimateur au sens des moindres carrés était :

$$\tilde{X} = \frac{\sum p_i X_i}{\sum p_i} \quad \text{avec} \quad \Gamma_{\tilde{X}} = \frac{\sigma_0^2}{\sum p_i} \quad (\sigma_0^2 \text{ variance unitaire}) \quad (3.3)$$

On peut sans peine étendre la validité de ces formules au cas où  $\tilde{X}$  est un vecteur.

**Théorème 3.1** *L'estimateur  $\tilde{X}$  au sens des moindres carrés d'une grandeur à  $r$  dimensions  $\tilde{X}$ , à partir d'un échantillon de  $N$  estimations indépendantes  $X_1, X_2, \dots, X_N$ , et sa variance  $\Gamma_{\tilde{X}}$  sont donnés par :*

$$\tilde{X} = \left( \sum_{i=1}^N P_i \right)^{-1} \sum_{i=1}^N P_i X_i = \frac{\sum_{i=1}^N P_i X_i}{\sum_{i=1}^N P_i}; \quad \Gamma_{\tilde{X}} = \frac{\sigma_0^2}{\sum_{i=1}^N P_i} \quad (3.4)$$

On peut suivre pour déterminer  $\tilde{X}$  une démarche absolument analogue à celle qui nous a servi à estimer un ensemble de  $r$  grandeurs à partir de  $n$  observations; on est conduit ici à minimiser la forme quadratique :

$$U = \sum_{i=1}^N (X_i - \tilde{X})^T P_i (X_i - \tilde{X})$$

On obtient, en exprimant que  $\partial U / \partial X = 0$  :

$$\sum_{i=1}^N P_i \tilde{X} = \sum_{i=1}^N P_i X_i$$

soit :

$$\tilde{X} = \left( \sum_{i=1}^N P_i \right)^{-1} \sum_{i=1}^N P_i X_i$$

De plus :

$$\begin{aligned} \Gamma_{\tilde{X}} &= E(\tilde{X} - E(\tilde{X})) \cdot (\tilde{X} - E(\tilde{X}))^T = \\ &= \left[ (\sum_i P_i)^{-1} \sum_i P_i E(X_i - E(X_i)) \right] \left[ (\sum_i P_i)^{-1} \sum_i P_i E(X_i - E(X_i)) \right]^T \\ \Gamma_{\tilde{X}} &= \left( \sum_i P_i \right)^{-1} \left[ \sum_i P_i E(X_i - E(X_i)) \right] \left[ \sum_i P_i E(X_i - E(X_i)) \right]^T \left( \sum_i P_i \right)^{-1} \end{aligned}$$

et comme  $E((X_i - E(X_i))(X_j - E(X_j)))^T = 0$  pour tout  $i \neq j$  (car les estimations  $X_1, X_2, \dots, X_N$  sont indépendants) :

$$\Gamma_{\tilde{X}} = \left( \sum_i P_i \right)^{-1} \left( \sum_i P_i \Gamma_{\tilde{X}_i} P_i \right) \left( \sum_i P_i \right)^{-1} = \left( \sum_i P_i \right)^{-1} \sigma_0^2$$

### 3.3 CONTRIBUTIONS DE CHAQUE EQUATION D'OBSERVATIONS, OU D'UNE PARTIE DU SYSTÈME DES EQUATIONS D'OBSERVATIONS, À LA MATRICE NORMALE DANS LE CAS OÙ LES OBSERVATIONS SONT INDÉPENDANTES

Soit un système d'équations d'observations :

$${}_n A_{rr} \dot{X}_1 - {}_n K_1 = {}_n \dot{V}_1 \quad (3.5)$$

La matrice normale correspondante est  $N = A^T P A$  ( $P$  diagonale); désignons par  $A_i$  le  $i$ ème vecteur ligne de  $A$  et par  $K_i, \dot{V}_i$  les  $i$ èmes composantes de  $K$  et  $V$ . Le système (3.5) s'écrit :

$$A_i \dot{X} - K_i = \dot{V}_i \quad (i = 1, n)$$

Et on a :

$$N = (A_1^T, A_2^T, \dots, A_n^T) \begin{pmatrix} p_1 & & & \\ & p_2 & & \\ & & \ddots & \\ & & & p_n \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix} = \sum_{i=1}^n p_i A_i^T A_i \quad (3.6)$$

$$A^T P K = \sum_{i=1}^n p_i A_i^T K_i \quad (3.7)$$

La contribution de la  $i$ ème équation d'observations au système  $NX - W = 0$  est donc :

$${}_r N_{i_r} = p_i A_i^T A_i \quad \text{et} \quad {}_r W_{i_r} \text{ (terme constant) } = p_i A_i^T K_i \quad (3.8)$$

$$\text{Et on a : } N = \sum N_i, \quad W = \sum N_i$$

On voit donc que la normalisation peut se faire pas à pas; ceci peut être intéressant en programmation au cas où on a un très grand nombre d'équations d'observations : on lit une équation, on calcule sa contribution qu'on ajoute à celle des équations déjà lues; il n'est pas besoin de garder en mémoire la matrice  $A$ .

On généralise immédiatement ces résultats à un groupe d'équations d'observations.

### 3.4 PROBLÈME DES OBSERVATIONS RÉPÉTÉES- CHOIX DES POIDS POUR UN SYSTÈME DE MESURES ENTIÈREMENT DÉTERMINÉ- CHOIX DES POIDS POUR DES OBSERVATIONS SUPPLÉMENTAIRES D'UNE GRANDEUR DÉJÀ OBSERVÉE

#### 3.4.1 Groupement des observations relatives à une même grandeur observée

Dans l'introduction nous avons supposé qu'il n'y avait qu'une mesure par grandeur observée; étant entendu que cette mesure était introduite avec son poids évalué en fonction de l'exactitude de la mesure.

Nous revenons ici sur cette question essentiellement à cause du problème de l'adjonction après coup de relations d'observations supplémentaires relatives à une grandeur, et accessoirement pour le problème de l'étude d'un système de mesures entièrement déterminé.

On suppose donc qu'on possède  $q$  mesures indépendantes  $L_{i1}, L_{i2}, \dots, L_{iq}$  d'une même grandeur observée  $\dot{L}_i$ , de poids  $p_1, p_2, \dots, p_q$ ; on va chercher alors la condition qui doit être remplie pour qu'il soit équivalent de poser les  $q$  équations d'observations, et de poser une seule équation relative à l'observation résumée  $\frac{\sum_j p_j L_{ij}}{\sum_j p_j}$ .

Toutes les équations d'observations relatives à une même grandeur observée  $\dot{L}_i$ , dérivent de la même relation géométrique :

$$\varphi(\dot{L}_i, \mathcal{X}_1) = 0 \quad (3.9)$$

qui par linéarisation donne (§ 2.2-formule (2.4)) :

$$a_{i1}\dot{X}_1 + a_{i2}\dot{X}_2 + \dots + a_{ir}\dot{X}_r + h_{i0} + L_{i0} - \dot{L}_i = 0 \quad (3.10)$$

Les constantes  $a_{i1} \dots a_{ir}$  et  $h_{i0}$  ne dépendent que des valeurs approchées de  $\dot{L}$  et  $\mathcal{X}$ , soit  $L_{i0}$  et  $\mathcal{X}_0$ .

D'où pour l'observation  $L_{ij}$ , l'équation d'observations :

$$a_{i1}\dot{X}_1 + a_{i2}\dot{X}_2 + \dots + a_{ir}\dot{X}_r + h_{i0} + L_{i0} - L_{ij} = \dot{V}_{ij} \quad (3.11)$$

On constate donc que tous les coefficients  $a_{i1}, \dots, a_{ir}$ , ainsi que la constante  $h_{i0} + L_{i0}$  sont indépendantes de l'observation  $L_{ij}$ .

Ceci étant, les  $q$  équations d'observations relatives à la grandeur observée  $\dot{L}_i$ , s'écrivent :

$$\begin{aligned} a_{i1}\dot{X}_1 + a_{i2}\dot{X}_2 + \dots + a_{ir}\dot{X}_r + h_{i0} + L_{i0} - L_{i1} &= \dot{V}_{i1} \\ a_{i1}\dot{X}_1 + a_{i2}\dot{X}_2 + \dots + a_{ir}\dot{X}_r + h_{i0} + L_{i0} - L_{i2} &= \dot{V}_{i2} \\ &\vdots \\ a_{i1}\dot{X}_1 + a_{i2}\dot{X}_2 + \dots + a_{ir}\dot{X}_r + h_{i0} + L_{i0} - L_{iq} &= \dot{V}_{iq} \end{aligned} \quad (3.12)$$

Soit alors  $\Delta N$  (matrice normale) et  $\Delta W$  (terme constant) la contribution de ces  $q$  équations au système normal; on a (paragraphe précédent) :

$$\Delta_r N_r = \sum_{j=1}^{j=q} p_j A_j^T A_j \text{ avec } A_j = (a_{j1}, a_{j2}, \dots, a_{jr})$$

$$\Delta_r W_1 = \sum_{j=1}^{j=q} p_j A_i^T (h_{i0} + L_{i0} - L_{ij})$$

ou encore :

$$\Delta N = \left( \sum_{j=1}^{j=q} p_j \right) A_i^T A_i, \quad \Delta W = \left( \sum_{j=1}^{j=q} p_j \right) A_i^T \left( h_{i0} + L_{i0} - \frac{\sum_{j=1}^{j=q} p_j L_{ij}}{\sum_{j=1}^{j=q} p_j} \right) \quad (3.13)$$

Considérons à présent l'observation "résumé" des  $q$  observations  $L_{i1}, L_{i2}, \dots, L_{iq}$  :

$$\tilde{L}_i = \frac{\sum_{j=1}^{j=q} p_j L_{ij}}{\sum_{j=1}^{j=q} p_j}; \quad \Gamma_{\tilde{L}_i} = \frac{\sigma_0^2}{\sum_{j=1}^{j=q} p_j} \quad (3.14)$$

elle donne lieu à l'unique équation d'observations :

$$a_{i1}\dot{X}_1 + a_{i2}\dot{X}_2 + \dots + a_{ir}\dot{X}_r + h_{i0} + L_{i0} - \frac{\sum_{j=1}^{j=q} p_j L_{ij}}{\sum_{j=1}^{j=q} p_j} = \dot{V}_i \quad (3.15)$$

Si nous prenons comme poids de  $\tilde{L}_i$  la quantité  $\left( \sum_{j=1}^{j=q} p_j \right)$  on voit alors que la contribution de cette unique équation d'observations au système normal est :

$$\Delta N' = \left( \sum_{j=1}^{j=q} p_j \right) A_i^T A_i = \Delta N$$

$$\Delta W' = \left( \sum_{j=1}^{j=q} p_j \right) A_i^T \left( h_{i0} + L_{i0} - \frac{\sum_{j=1}^{j=q} p_j L_{ij}}{\sum_{j=1}^{j=q} p_j} \right) = \Delta W$$

**Conclusion 3.1** *Etant donné  $q$  observations  $L_{i1}, L_{i2}, \dots, L_{iq}$  d'une grandeur observée, il est équivalent (mathématiquement) de poser les  $q$  équations d'observations correspondantes, et de poser une seule équation avec l'observation résumée  $\sum_{j=1}^{j=q} p_j L_{ij} / \sum_{j=1}^{j=q} p_j$  à condition que le poids de l'observation résumée soit la somme des poids des observations élémentaires.*

Malheureusement en pratique, cette équivalence n'est vraie que si on admet que "précision" et "exactitude" se confondent : dans ce cas le poids de l'observation résumée (inversement proportionnel à sa variance  $\sigma_0^2 / \sum_j p_j$ ) est effectivement la somme des poids des observations élémentaires.

Or nous avons vu (introduction) qu'il n'est pas raisonnable de faire l'hypothèse qu'il en est toujours ainsi : cela reviendrait à considérer que l'exactitude d'une observation résumée s'accroît à l'infini quand on multiplie les observations élémentaires. Aussi avons nous supposé qu'on n'introduisait par grandeur observée qu'une seule observation résumée avec un poids estimé en fonction de son exactitude.

On peut cependant envisager une exception : le cas où on répète les observations dans des conditions expérimentales identiques.

### 3.4.2 Etude d'un système de mesures entièrement déterminé

Par système de mesures entièrement déterminé nous entendons un système où les conditions d'observations des  $n$  grandeurs observées sont entièrement fixées (instruments, observateurs, méthodes,...). Dans ce cas la loi des mesures  $L_i$  est normale  $\mathcal{N}(E(L_i), \sigma_0 / \sqrt{p_i})$  avec  $E(L_i) \neq \dot{L}_i$  (puisque il y a toujours un biais), et  $p_i$  non pas inversement proportionnel au carré de l'erreur quadratique (exactitude) mais au carré de l'écart-type.

Si le but est alors d'étudier la stabilité (ou la "répétabilité") d'un tel système c'est-à-dire d'étudier la cohérence des estimations qu'on obtient en répétant la série des  $n$  mesures, on s'appuiera sur le fait que la loi des observations  $L_i$  relative à un tel système est à priori normale  $\mathcal{N}(E(L_i), \sigma_0 / \sqrt{p_i})$ .



Dans ce cas, il est absolument équivalent de poser une équation d'observations par observation, et d'en poser une seule pour l'observation résumée du moment qu'on calcule les poids à partir des précisions et non des exactitudes.

### 3.4.3 *Choix des poids d'observations supplémentaires relatives à une grandeur observée*

Nous avons supposé (à l'exception du cas du paragraphe précédent) que pour chacune des  $n$  grandeurs observées on n'introduisait qu'une seule mesure résumée avec un poids évalué en fonction de son exactitude.

Mais il peut se faire (raccordement des réseaux géodésiques, addition in extremis d'observations supplémentaires) qu'on ait après coup à tenir compte d'observations supplémentaires.

Le problème est d'évaluer correctement les poids de ces observations supplémentaires.

Soit  $L_{i(m)}$  l'observation résumée de  $m$  mesures élémentaires relatives à la grandeur observée  $L_i$  et  $p_{i(m)}$  son poids.

Soit  $L'_{i(m)}$  une observation supplémentaire de la même grandeur observée  $L_i$ .

Nous supposons connue une estimation de la justesse des mesures de la grandeur  $L_i$ , soit  $\sqrt{E(\beta^2)}$  ( $\beta$  étant le biais d'une mesure).

Si  $\sigma_0^2$  est la variance unitaire, et  $\sigma_{L_{i(m)}}$  la précision (écart-type) de l'observation résumée  $L_{i(m)}$ , on a par définition :

$$p_{i(m)} = \frac{\sigma_0^2}{emq_{L_{i(m)}}^2} = \frac{\sigma_0^2}{E(\beta^2) + \sigma_{L_{i(m)}}^2} \leq \frac{\sigma_0^2}{E(\beta^2)} \quad (3.16)$$

avec  $q_1, q_2, \dots, q_m$  sont les poids des  $m$  observations élémentaires dont le résumé est  $L_{i(m)}$  (§ chapitre 2) :

$$\sigma_{L_{i(m)}}^2 = \frac{\sigma^2}{\sum_1^m q_i} \quad (\sigma^2 \text{ variance unitaire correspondante}) \quad (3.17)$$

Ceci étant, voyons le poids qu'il aurait fallu affecter à la mesure résumée, si dès le départ, on avait tenu compte des  $(m + 1)$  mesures; c'est par définition :

$$p_{i(m+1)} = \frac{\sigma_0^2}{E(\beta^2) + \sigma_{L_{i(m+1)}}^2} \quad \text{avec} \quad \sigma_{L_{i(m+1)}}^2 = \frac{\sigma^2}{\sum_1^{m+1} q_i} \quad (3.18)$$

Il faut donc introduire l'observation supplémentaire  $L'_i$  avec le poids :

$$p'_i = p_{i(m+1)} - p_{i(m)}$$

En effet, la pose de deux équations d'observations, la première relative à  $L_{i(m)}$ , de poids  $p_{i(m)}$ , la seconde relative à  $L'_i$  de poids  $p'_i$ , équivaut formellement à la pose de l'unique observation résumée (3.4.1) :

$$\frac{p_{i(m)}L_{i(m)} + p'_iL'_i}{p_{i(m)} + p'_i} \quad \text{de poids} \quad p_{i(m)} + p'_i = p_{i(m+1)}$$

**Conclusion 3.2** *Le poids d'une observation relatif à une grandeur observée  $L_i$  est toujours inférieur à  $\frac{\sigma_0^2}{E(\beta^2)}$ ,  $\sqrt{E(\beta^2)}$  caractérisant l'exactitude maximale (justice) avec laquelle on peut déterminer cette grandeur dans tous les systèmes de mesures analogues.*

On évalue alors le poids d'une observation supplémentaire  $L'_i$ , par :

$$p'_i = \frac{\sigma_0^2}{emq_{L_{i(m)}}^2} - \frac{\sigma_0^2}{emq_{L_{i(m+1)}}^2}$$

avec :

- $emq_{L_{i(m)}} = emq$  de la moyenne pondérée des  $m$  mesures déjà introduites.
- $emq_{L_{i(m+1)}} = emq$  de la moyenne pondérée des  $m + 1$  mesures (la  $(m + 1)$ ème étant  $L'_i$ ).

On voit que les poids à introduire par observation supplémentaire sont dégressifs<sup>1</sup>; en pratique ils le sont rapidement : une deuxième mesure n'apporte pour ainsi dire rien aux trois premières (§ chapitre 1).

1. On rompt ainsi la symétrie du modèle classique de la théorie des erreurs, où toutes les observations sont introduites avec leur poids propre, et où la solution finale est indépendante de l'ordre dans lequel on introduit les mesures supplémentaires; mais cela est sans importance si on considère le fait que toutes les observations résumées de même exactitude sont équivalentes.

Notons que si on ne procède pas ainsi, on risque en rajoutant inconsidérément des observations supplémentaires de très surestimer le poids relatif à la grandeur observée correspondante.

### 3.5 ADJONCTION D'OBSERVATIONS OU D'INCONNUES SUPPLÉMENTAIRES (cas de grands réseaux) - MÉTHODE DU BORDAGE

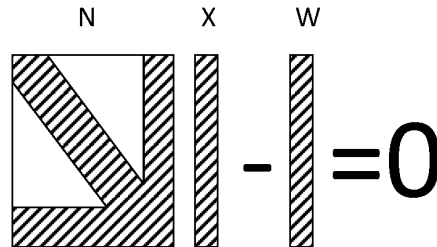
Dans les problèmes de compensation de réseaux (géodésie, photogrammétrie), on a généralement affaire, dès que le réseau atteint une certaine taille, et à condition que les équations d'observations et les inconnues soient rangées dans un certain ordre, à un système en forme de bande à largeur étroite (le long de la diagonale principale) avec une bordure à droite.



Une telle structure est possible pour des raisons topologiques : dans un réseau un point n'est généralement relié aux autres que par l'intermédiaire de mesures relatives à quelques points immédiatement voisins ; la bordure de droite du système regroupe les inconnues communes à toutes ou à un grand nombre d'équations d'observations (par exemple inconnues rendant compte de systématismes).

Comme on le voit facilement, cette structure se conserve après normalisation (à ceci près qu'il y a maintenant deux marges symétriques par rapport à la diagonale principale) ; le système normal à l'allure suivante (Fig. 3.1) :

$$NX - W = 0 \quad (3.19)$$



**Fig. 3.1** Allure du système normal

La remarque qu'une telle structure est possible est importante, car pour les systèmes d'ordre élevé, ce type de structure est très favorable aux techniques d'inversion (à condition bien sûr que les bordures soient relativement étroites).

Le problème que nous nous proposons d'examiner est alors le suivant :

- un processus de programmation a été mis au point, qui aboutit à partir d'un certain nombre d'équations d'observations, à un système normal du type (3.19).
- on veut alors sans reprendre la programmation de cette partie (souvent très complexe) rajouter (ou retrancher) certaines équations d'observations, et éventuellement, dans le cas où on rajoute des équations d'observations pouvoir introduire des inconnues supplémentaires, comment opérer?

Désignons par :  ${}_n A_{r,r} \dot{X}_1 - {}_n K_1 = {}_n \dot{V}_1$

le système des équations d'observations initiales de matrice de poids  $P$ , et par :

$${}_p B_{r,r} \dot{X}_1 - {}_p M_1 = {}_p \dot{V}'_1$$

les  $p$  nouvelles équations d'observations (nous supposons d'abord qu'il n'y a pas d'inconnues supplémentaires) de matrice de poids  $Q$ .

Le système regroupant les  $(n + p)$  équations d'observations est :

$$\begin{pmatrix} A \\ B \end{pmatrix} \cdot_r \dot{X}_1 - \begin{pmatrix} K \\ M \end{pmatrix} = \begin{pmatrix} \dot{V} \\ \dot{V}' \end{pmatrix}$$

avec la matrice de poids  $(n + p) \times (n + p)$  :

$$\begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix}$$

Le système normal correspondant est :

$$(N + B^T Q B).X - (W + B^T Q M) = 0 \quad (3.20)$$

On voit très facilement qu'il est équivalent, au système suivant :

$$\begin{pmatrix} N & B^T \\ B & -Q^T \end{pmatrix} \cdot \begin{pmatrix} X \\ \Lambda \end{pmatrix} - \begin{pmatrix} W \\ M \end{pmatrix} = 0 \quad (3.21)$$

l'élimination du corrélatif  ${}_p\Lambda_1$  redonne en effet (3.20).

**Conclusion 3.3** *Etant donné un système normal  $NX - W = 0$ , on peut sans modifier sa structure (bande +bordure) rajouter de nouvelles équations d'observations  $B\tilde{X} - M = \tilde{V}$ , de poids  $Q$  en bordant  $N$  à droite par  $B^T$ ,  $(-Q)^{-1}$ , en bas par  $B$ , en allongeant le vecteur  $X$  du corrélatif  $\Lambda$  et le vecteur constant  $W$ , de  $M$  :*

$$\left. \begin{array}{l} {}_r N_r \cdot {}_r X_1 - {}_r W_1 = 0 \\ {}_p B_r \cdot {}_r X_1 - {}_p M_1 = {}_p \tilde{V}'_1 \\ \text{Poids : } {}_p Q_p \end{array} \right\} \Rightarrow \begin{pmatrix} {}_r N_r & {}_r B_p^T \\ {}_p B_r & -{}_p Q_p^T \end{pmatrix} \cdot \begin{pmatrix} {}_r \tilde{X}_1 \\ {}_p \Lambda_1 \end{pmatrix} - \begin{pmatrix} {}_r W_1 \\ {}_p M_1 \end{pmatrix} = 0 \quad (3.22)$$

$\tilde{X}$  solution au sens des moindres carrés de l'ensemble des équations d'observations

**Remarques :**

a) **Observations de poids infini**

Dans ce cas  $Q^{-1} = 0$ , le système (3.22) prend alors une forme que nous connaissons déjà (système normal correspondant à un système d'équations d'observations avec équations de condition liant les inconnues).

b) **Elimination d'équations d'observations**

On les récrit sous forme d'équations d'observations supplémentaires, avec le poids  $(-Q)$  (au lieu de  $+Q$ ).

c) **Equations d'observations supplémentaires avec inconnues supplémentaires**

Si on introduit les nouvelles inconnues  $\dot{x}$ , on pose :

$${}^p Y_1 = \begin{pmatrix} \dot{X} \\ \dot{x} \end{pmatrix}$$

Les équations d'observations supplémentaires s'écrivent alors :

$${}^p B {}^p Y_1 - M = \dot{V}_S \quad (3.23)$$

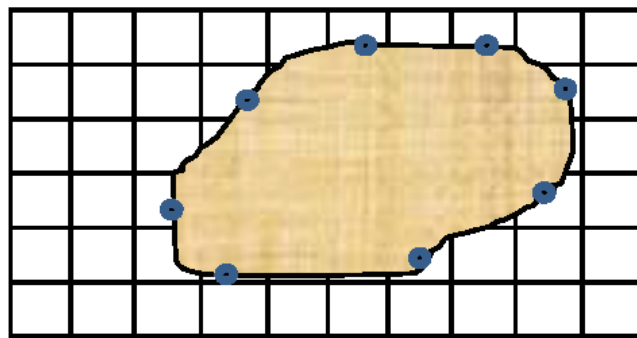
alors que les équations d'observations initiales (qui ont conduit au système normal  $NX - W = 0$ ) peuvent s'écrire :

$$(A, 0) \dot{Y} - \begin{pmatrix} K \\ 0 \end{pmatrix} = \begin{pmatrix} \dot{V} \\ 0 \end{pmatrix}$$

On constate immédiatement que le fait de rajouter les  $p$  équations supplémentaires (3.23), conduit à modifier le système  $NX - W = 0$  de la façon suivante :

$$\begin{pmatrix} N & 0 & & \\ 0 & 0 & & B^T \\ \hline & & B & -Q^{-1} \end{pmatrix} \cdot \begin{pmatrix} {}^p Y_1 \\ {}^p \Lambda_1 \end{pmatrix} - \begin{pmatrix} W \\ 0 \\ M \end{pmatrix} = 0 \quad (3.24)$$

Exemple : le problème du lac (ou des lacs) d'altitude inconnue (aéro-triangulation analytique) : Soit  $z_0$  l'altitude du lac, inconnue; on peut désirer pour amélio-



**Fig. 3.2** Le problème du lac

rer l'exactitude en altéométrie, écrire qu'un certain nombre de points  $i$  le long

du lac (voire au milieu) ont pour altitude  $z_0$ , c'est-à-dire écrire des équations d'observations du type :

$$dz_i - (z_0 - z_i) = v_i$$

Si on n'a pas prévu ce type d'équations, (il y a une inconnue d'altitude  $z_0$  par lac...), il est commode d'en tenir compte après coup par la méthode du bordage.





Chapitre **4**

# Propriétés Statistiques des Estimateurs des Moindres Carrés

## Sommaire

---

|            |  |           |
|------------|--|-----------|
| <b>4.1</b> | <b>INTRODUCTION</b> .....  | <b>75</b> |
| <b>4.2</b> | <b>PROPRIÉTÉS STATISTIQUES DES ESTIMATEURS DES MOINDRES CARRÉS DANS L'HYPOTHÈSE OÙ LES ERREURS DE MESURES SONT ACCIDENTELLES- VARIANCE DES ESTIMATEURS</b>   | <b>76</b> |
| 4.2.1      | Les estimateurs des moindres carrés .....  | 76        |
| 4.2.2      | Estimateur de la variance unitaire $\sigma_0^2$ .....  | 77        |
| 4.2.3      | Expression des matrices-variance des inconnues, des observations compensées, des résidus, de toute fonction linéaire des inconnues .....                     | 80        |
| <b>4.3</b> | <b>PROPRIÉTÉS STATISTIQUES DES ESTIMATEURS DES MOINDRES CARRÉS DANS L'HYPOTHÈSE OÙ LES ERREURS DE MESURES SONT NORMALES ET CENTRÉES</b> .....                | <b>85</b> |
| <b>4.4</b> | <b>COMPARAISON DE LA VARIANCE UNITAIRE A PRIORI <math>\sigma_0^2</math> QUAND ELLE EST CONNUE AVEC LA VARIANCE UNITAIRE ESTIMÉE <math>s_0^2</math></b> ..... | <b>93</b> |

---

#### 4.1 INTRODUCTION

Lorsque nous avons cherché à estimer un ensemble de  $r$  grandeurs inconnues  ${}_r\mathcal{X}_1$  à partir de  $n$  observations  ${}_nL_1$  que nous avons supposées correspondre toutes à des grandeurs observées distinctes, nous avons d'abord posé que le vecteur  $L$  était normal  $\mathcal{N}(\dot{L}, \sqrt{P^{-1}}\sigma_0)$ ,  $\sigma_0$  étant la variance de la mesure de poids unitaire, et  $P$  la matrice des poids.

La règle du maximum de vraisemblance conduit alors pour estimer au mieux les grandeurs inconnues à minimiser la forme quadratique fondamentale :

$$\dot{V}^T P \dot{V} \quad \text{avec} \quad \dot{V} = \dot{L} - L$$

ce qui, compte-tenu des relations entre les inconnues et les résidus vrais  $\dot{V}$  donne pour  $\dot{X}, \dot{L}, \dot{V}$  et  $\sigma_0^2$  des estimateurs que nous avons désignés par :

$$\tilde{X}, \tilde{L}, \tilde{V} \quad \text{et} \quad s_0^2 = \frac{1}{n} \tilde{V}^T P \tilde{V}$$

Ces estimateurs (conséquences de l'emploi de la règle du maximum de vraisemblance) sont asymptotiquement (au moins) sans biais, normaux, et "efficaces".

Dans ce qui suit nous supposerons d'abord que les erreurs sont seulement accidentelles ou sans biais :  $E(\dot{L} - L) = 0$ , et non pas nécessairement normales, et allons étudier sous couvert de cette seule hypothèse les propriétés des estimateurs des moindres carrés. Nous compléterons ensuite par l'énoncé des propriétés supplémentaires découlant de l'hypothèse plus forte de normalité des observations.

## 4.2 PROPRIÉTÉS STATISTIQUES DES ESTIMATEURS DES MOINDRES CARRÉS DANS L'HYPOTHÈSE OÙ LES ERREURS DE MESURES SONT ACCIDENTELLES- VARIANCE DES ESTIMATEURS

### 4.2.1 Les estimateurs des moindres carrés

**Théorème 4.1.** *Les estimateurs des moindres carrés des inconnues, des observations, des résidus vrais et de toute fonction linéaire des inconnues sont sans biais.*

*Autrement dit :*

$$E(\dot{L} - L) = 0 \implies \begin{cases} E(\tilde{X}) = \dot{X} \\ E(\dot{L}) = \dot{L} \\ E(\dot{V}) = 0 \\ E(D\tilde{X} + F) = D\dot{X} + F \end{cases} \quad (4.1)$$

Bien entendu, s'il y a eu au moment de la pose des équations d'observations ou de condition, linéarisation, ces propriétés ne sont vraies que dans la mesure où les termes négligés sont suffisamment petits.

\* Démontrons d'abord que  $E(\tilde{X}) = \dot{X}$  :

On peut toujours supposer puisqu'un estimateur au sens des moindres carrés est unique que l'estimateur  $\tilde{X}$  est déterminé par exemple à partir d'un système d'équations d'observations :

$$A\tilde{X} - (H + L) = \dot{V} \quad \text{avec} \quad E(\dot{V}) = 0 \quad (4.2)$$

avec  $H$  ne dépendant pas de  $L$ , mais uniquement des valeurs approchées des inconnues ( voir (2.4)), or la solution des moindres carrés est :

$$\tilde{X} = N^{-1} .A^T P(H + L)$$

d'où :  $E(\tilde{X}) = N^{-1} .A^T P E(H + L) = N^{-1} .A^T P(H + \dot{L}) = \dot{X}$ , car (4.2) entraîne  $\dot{X} = N^{-1} .A^T P(H + \dot{L})$ , donc  $\tilde{X}$  est non biaisé.

\* Il est de même pour toute fonction linéaire des inconnues, en effet soit :

$$\dot{Y} = D\tilde{X} + F$$

On sait que l'estimateur  $\tilde{Y}$  des moindres carrés est :  $\tilde{Y} = D\tilde{X} + F$ , on a alors :

$$E(\tilde{Y}) = E(D\tilde{X} + F) = D.E(\tilde{X}) + F = D\hat{X} + F = \hat{Y}$$

\* D'où il suit, puisque  $\tilde{L} = A\tilde{X} - H \implies E(\tilde{L}) = \hat{L}$ .

\* Enfin,  $E(\tilde{V}) = E(\tilde{L} - L) = 0$ .

#### 4.2.2 Estimateur de la variance unitaire $\sigma_0^2$

**Théorème 4.1** *Un estimateur sans biais de  $\sigma_0^2$  est :*

$$s_0^2 = \frac{1}{n-r} \tilde{V}^T . P . \tilde{V}$$

(4.3)

*n nombre de grandeurs observées  
r nombre de degrés de liberté des inconnues*

Là aussi, à cause de l'importance de la solution des moindres carrés, on peut utiliser pour établir ce résultat, l'expression obtenue à partir d'un système d'équations d'observations.

Partant de :

$$A\hat{X} - H = L + \hat{V} \quad (4.4)$$

$$A\tilde{X} - H = L + \tilde{V} \quad (4.5)$$

Il vient d'abord (4.4) :

$$\hat{X} = N^{-1}A^T P(H + L) + N^{-1}A^T P\hat{V} = \tilde{X} + N^{-1}A^T P\hat{V}$$

D'où on déduit :

$$A(\hat{X} - \tilde{X}) = AN^{-1}A^T P\hat{V} \quad (4.6)$$

Mais de (4.4) et (4.5), on tire aussi :

$$A(\hat{X} - \tilde{X}) = \hat{V} - \tilde{V} \quad (4.7)$$

Soit, rapprochant (4.7) de (4.6) :

$$\tilde{V} = (I - AN^{-1}A^T P)\hat{V} \quad (4.8)$$

Formule importante, où on voit les résidus  $\tilde{V}$  exprimés en fonction des résidus vrais  $\dot{V}$  (inconnus). La formule n'est évidemment valable que s'il n'y a pas d'équations de condition.

De (4.8), on déduit immédiatement :

$$\sqrt{P}\tilde{V} = (I - \sqrt{P}AN^{-1}A^T\sqrt{P})\sqrt{P}\dot{V}$$

ou encore :

$$\boxed{\begin{array}{l} \sqrt{P}\tilde{V} = U\sqrt{P}\dot{V} \\ \text{avec } U = I - S \\ S = \sqrt{P}AN^{-1}A^T\sqrt{P} \end{array}} \quad (4.9)$$

On vérifie sans difficulté que  $U$  et  $S$  sont symétriques et idempotentes ( $S^n = S$  et  $U^n = U$ ). Cherchons alors la variance du vecteur  $\sqrt{P}\tilde{V}$ , on a :

$$\Gamma_{\sqrt{P}\tilde{V}} = U\Gamma_{\sqrt{P}\dot{V}}U^T$$

où  $\Gamma_{\sqrt{P}\dot{V}}U^T$  est la variance du vecteur  $\sqrt{P}\dot{V}$  qui vaut  $\sigma_0^2 I$  (puisque  $\Gamma_{\dot{V}} = P^{-1} \cdot \sigma_0^2$  par définition).

Par suite :

$$\Gamma_{\sqrt{P}\tilde{V}} = UU^T\sigma_0^2 = U\sigma_0^2$$

$$\boxed{\Gamma_{\sqrt{P}\tilde{V}} = (I - \sqrt{P}AN^{-1}A^T\sqrt{P})\sigma_0^2} \quad (4.10)$$

(pas d'équations de condition liant les inconnues)

On peut à présent calculer l'espérance de  $\tilde{V}^T P \tilde{V}$ , carré de la norme du vecteur  $\sqrt{P}\tilde{V}$ ; c'est, puisque  $\sqrt{P}\tilde{V}$  est sans biais, la somme des variances marginales de  $\Gamma_{\sqrt{P}\tilde{V}}$  c'est-à-dire la trace de la matrice  $\Gamma_{\sqrt{P}\tilde{V}}$  :

$$E(\tilde{V}^T P \tilde{V}) = Tr(\Gamma_{\sqrt{P}\tilde{V}}) = Tr(U)\sigma_0^2 \quad (4.11)$$

La trace de  $U$  somme des éléments de la diagonale principale, est aussi comme on sait, somme des valeurs propres de  $U$ ; nous allons montrer qu'elle vaut  $(n - r)$ . Pour cela considérons d'abord la matrice :

$${}_n S_n = {}_n \sqrt{P}_{n \cdot n} A_{r r} N_r^{-1} {}_r A_n^T {}_n \sqrt{P}_n$$

Il clair que l'on a, puisque  $N^{-1}$  et  $A$  sont de rang  $r$  (et vues les dimensions de  $N^{-1}$  et  $A$ ) :  $\text{rang } S = r$ .

$S$  a donc  $(n - r)$  valeurs propres nulles, et  $r$  valeurs propres non nulles, qui sont toutes égales à l'unité car  $S$  est idempotente (en effet si  $\lambda$  est une valeur propre de  $S$  et  $X$  le vecteur propre correspondant,  $SX = \lambda X \implies S^2X = \lambda^2X = \lambda X \implies \lambda = 0$  ou  $+1$ ; signe  $+$  car  $X^T SX \geq 0$ ).

Il s'ensuit que la matrice  $U = I - S$ , qui est telle que  $US = 0$ , a  $r$  valeurs propres nulles et les  $(n - r)$  autres égales à l'unité; en effet soit  $V_j (j = 1, r)$  les vecteurs propres de  $S$  correspondant aux valeurs propres de  $S$  égales à 1 :

$$UV_j = V_j - SV_j = 0$$

tous les  $V_j$  sont donc valeurs propres de  $U$  avec la valeur propre 0; de même, si  $W_j (j = 1, n - r)$  sont les vecteurs propres de  $S$  correspondant aux valeurs propres nulles de cette matrice, on a :

$$UW_j = W_j - SW_j = W_j$$

$W_j$  est vecteur propre avec la valeur propre 1.  $U$  a donc  $r$  valeurs propres nulles, et  $(n - r)$  valeurs propres égales à 1.

Par suite :  $\text{Tr}(U) = n - r$  et de (4.11)  $E(\tilde{V}^T P \tilde{V}) = (n - r)\sigma_0^2$ , soit :

$$\sigma_0^2 = \frac{E(\tilde{V}^T P \tilde{V})}{n - r} \quad (4.12)$$

D'où on déduit un estimateur sans biais de  $\sigma_0^2$  :

$$s_0^2 = \frac{\tilde{V}^T P \tilde{V}}{n - r} \quad (4.13)$$

car on a bien :

$$E(s_0^2) = \sigma_0^2$$

On notera que  $r$  est le nombre de degrés de liberté des inconnues, c'est-à-dire s'il y a des équations de condition liant les inconnues, le nombre d'inconnues moins le nombre d'équations de condition.

On notera aussi que l'estimateur  $s_0'^2$  de  $\sigma_0^2$  fourni directement (2.37) par l'application de la règle du maximum de vraisemblance, n'est pas sans biais, mais seulement asymptotiquement sans biais :

$$s_0'^2 = \frac{1}{n} \tilde{V}^T P \tilde{V} \implies E(s_0'^2) \neq \sigma_0^2$$

### 4.2.3 Expression des matrices-variance des inconnues, des observations compensées, des résidus, de toute fonction linéaire des inconnues

#### 4.2.3.1 Cas où les inconnues sont estimées à partir d'un système d'équations d'observations $A\tilde{X} - H = \tilde{L}$ sans équations de condition liant les inconnues

a) Variance de  $\tilde{X}$  (ou de  $\tilde{\mathcal{X}} = \mathcal{X}_0 + \tilde{X}$ )

L'estimateur  $\tilde{X}$  est :  $\tilde{X} = N^{-1}A^T P(H + L)$ . On a donc :

$$\Gamma_{\tilde{X}} = E((\tilde{X} - E(\tilde{X}))(\tilde{X} - E(\tilde{X}))^T) = N^{-1}A^T P \Gamma_L (N^{-1}A^T P)^{-1}$$

avec  $\Gamma_L = P^{-1}\sigma_0^2$ , soit :

$$\boxed{\Gamma_{\tilde{X}} = N^{-1}\sigma_0^2} \quad (4.14)$$

formule généralisant celle qui donne la variance de la moyenne de  $n$  observations d'écart-type  $\sigma : \sigma_0^2$ .

Si on ne connaît pas la variance unitaire  $\sigma_0^2$  on lui substituera son estimation  $s_0^2$ ; on obtiendra alors l'estimation sans biais de la variance des inconnues :

$$\tilde{\Gamma}_{\tilde{X}} = N^{-1}s_0^2 \quad \text{avec} \quad s_0^2 = \frac{1}{n-r} \tilde{V}^T P \tilde{V} \quad (4.15)$$

On notera qu'en général les inconnues estimées sont corrélées :  $N^{-1}$  n'est pas en général diagonale; on notera également que si on multiplie convenablement le nombre de grandeurs observées, on fera tendre vers 0 la matrice  $\Gamma_{\tilde{X}}$ .

#### Application à l'exemple fondamental

On a trouvé (chapitre 2, § 2.5.4.6) :  $s_0^2 = 0.68$

Mais comme les *emq* relatives aux diverses mesures ont été ici déterminées à priori avec un certain soin, on prendra plutôt pour valeur de la variance unitaire celle de la variance unitaire à priori, et on utilisera la formule (4.15) : Soit (chapitre 2, § 2.5.4.6) :

$$\begin{pmatrix} 2.06 & 2.24 & 1.08 \\ 2.24 & 3.01 & 1.86 \\ 1.08 & 1.86 & 2.12 \end{pmatrix} \quad (4.16)$$

b) Variance d'une fonction linéaire des inconnues :  $\dot{Y} = D\dot{X} + F$

$\tilde{X}$  est l'estimateur de  $\dot{X}$  que donne l'application de la méthode des moindres carrés, on sait que celui de  $Y$  est :

$$\tilde{Y} = D\tilde{X} + F$$

D'où on déduit :

$$\Gamma_{\tilde{Y}} = D\Gamma_{\tilde{X}}D^T \quad (4.17)$$

c) Variance des observations compensées  $\tilde{L}$  et des résidus  $\tilde{V}$

De (4.17) et de l'expression de  $\dot{L}$  :  $\dot{L} = A\dot{X} - H$ . On déduit :  $\Gamma_{\dot{L}} = A\Gamma_{\dot{X}}A^T$ .

Soit :

$$\Gamma_{\dot{L}} = AN^{-1}A^T\sigma_0^2 \quad (4.18)$$

Pour calculer la variance de  $\tilde{V}$ , on partira de la formule (4.8) :

$$\tilde{V} = (I - AN^{-1}A^T P)\dot{V}$$

D'où, en remarquant que  $\Gamma_{\dot{V}} = P^{-1}\sigma_0^2$  (puisque  $\dot{V} = \dot{L} - L$ ) :

$$\Gamma_{\tilde{V}} = (I - AN^{-1}A^T P)P^{-1}(I - AN^{-1}A^T P)\sigma_0^2$$



Soit tous calculs faits :

$$\Gamma_{\tilde{V}} = (I - AN^{-1}A^T)\sigma_0^2 \quad (4.19)$$

On notera que les observations compensées sont généralement corrélées, et qu'il en va de même pour les résidus. En particulier il en est généralement ainsi, lorsque les observations initiales  $L$  sont indépendantes, c'est-à-dire lorsque  $P$  est diagonale. Enfin, on voit que les vecteurs  $\tilde{L}$  et  $\tilde{V}$  sont non corrélés; cette propriété est évidemment indépendante de la méthode de pose des équations.

En effet, tenant compte de (4.8) et du fait que  $E(\tilde{V}) = 0$ , il vient :

$$E((\tilde{L} - E(\tilde{L}))(\tilde{V} - E(\tilde{V}))^T) = E((A\tilde{X} - A\hat{X})\hat{V}^T(I - AN^{-1}A^T P))$$

soit :

$$\begin{aligned} E((\tilde{L} - E(\tilde{L}))(\tilde{V} - E(\tilde{V}))^T) &= E(AN^{-1}A^T P(L - \hat{L})V^T(I - AN^{-1}A^T P)) \\ &= -AN^{-1}A^T P(P^{-1}\sigma_0^2)(I - AN^{-1}A^T P)^T \end{aligned}$$

(puisque  $E((L - \hat{L})\hat{V}^T) = -P^{-1}\sigma_0^2$ ), soit :

$$E((\tilde{L} - E(\tilde{L}))(\tilde{V} - E(\tilde{V}))^T) = 0$$

La covariance de  $\tilde{L}$  et  $\tilde{V}$  est nulle;  $\tilde{L}$  et  $\tilde{V}$  sont donc non corrélés; si de plus  $\tilde{L}$  et  $\tilde{V}$  sont normaux (nous verrons que c'est le cas si les observations sont normales), ils sont alors indépendants.

#### 4.2.3.2 Cas où les inconnues sont estimées à partir d'un système d'équations d'observations, avec équations de condition liant les inconnues : estimation de la variance des inconnues

Le système des moindres carrés a la forme suivante :

$$\begin{cases} {}_n A_{rr} \hat{X}_1 - K = \hat{V} \\ {}_p B_{rr} \hat{X}_1 - M = 0 \\ \text{avec } K = H + L \\ H \text{ et } M \text{ ne dependent pas de } L \\ \text{rang}(A) = r, \text{ rang}(B) = p \end{cases} \quad (4.20)$$

Dans ce cas nous avons vu que l'estimateur des moindres carrés  $\tilde{X}$  est solution de :

$$\begin{pmatrix} A^T PA & B^T \\ B & 0 \end{pmatrix} \cdot \begin{pmatrix} X \\ \Lambda \end{pmatrix} = \begin{pmatrix} A^T PK \\ M \end{pmatrix} \quad (4.21)$$

( $\Lambda$  corrélatif de Lagrange).

Nous nous proposons d'estimer  $\Gamma_{\tilde{X}}$ . Posons comme dans le cas du paragraphe précédent :  $N = A^T PA$ .

A partir de (4.21), on peut calculer  $\tilde{X}_i$  on trouve sans difficulté :

$$\tilde{X} = N^{-1}[I - B^T(BN^{-1}B^T)^{-1}BN^{-1}]A^T PK + N^{-1}B^T(BN^{-1}N^T)^{-1}M \quad (4.22)$$

Remarquons alors que seul le premier terme du second membre dépend des observations  $L$  (par l'intermédiaire de  $K$ ) ; il s'ensuit :

$$\Gamma_{\tilde{X}} = N^{-1}[I - B^T(BN^{-1}B^T)^{-1}BN^{-1}]A^T P\Gamma_L PA[I - N^{-1}B^T(BN^{-1}B^T)^{-1}B]N^{-1}$$

soit, tenant compte du fait que  $\Gamma_L = P^{-1}\sigma_0^2$  :

$$\Gamma_{\tilde{X}} = N^{-1}[N - B^T(BN^{-1}B^T)^{-1}B][I - N^{-1}B^T(BN^{-1}B^T)^{-1}B]N^{-1}\sigma_0^2$$

ou encore :

$$\Gamma_{\tilde{X}} = N^{-1}[I - B^T(BN^{-1}B^T)^{-1}BN^{-1}]\sigma_0^2 \quad (4.23)$$

Rapprochant de (4.22), on voit qu'au facteur  $\sigma_0^2$  près,  $\Gamma_{\tilde{X}}$  est la sous-matrice carrée d'ordre  $r$  supérieure gauche de la matrice inverse du système (4.21) ce que nous symboliserons par :

$$\Gamma_{\tilde{X}} = {}_r(\delta_{ij})_{r+p} \begin{pmatrix} A^T PA & B^T \\ B & 0 \end{pmatrix}_{(r+p) \times (r+p)}^{-1} \cdot {}_{r+p}(\delta_{ij})_r \sigma_0^2 \quad (4.24)$$

$\delta_{ij} = 1$  si  $i = j \leq r$ ; 0 dans les autres cas

Bien entendu, on pourrait ensuite donner à partir de cette expression celle de  $\Gamma_L, \Gamma_{\tilde{Y}}$  ou d'une fonction linéaire des inconnues.

### Application Numérique

Prenons l'énoncé de l'exemple fondamental, avec la condition exprimant que le triangle est rectangle (chapitre 2, § 2.2.1.2). Les calculs ont été effectués en (chapitre 2, § 2.5.4.6) ; la matrice variance  $\Gamma_{\tilde{X}}$  s'en extrait immédiatement ( $\sigma_0^2 = 1$ ) :

$$\Gamma_{\bar{X}} = \begin{pmatrix} 2.04 & 2.27 & 1.03 \\ 2.27 & 2.96 & 1.94 \\ 1.03 & 1.94 & 1.96 \end{pmatrix}$$

#### 4.2.3.3 Cas où on utilise la méthode des équations de condition liant les observations

Dans ce cas, on est amené à résoudre par moindres carrés, comme on sait, un système de la forme :

$$\begin{aligned} \dot{Y} &= \dot{V} \\ C\dot{Y} - Q &= 0 \end{aligned} \quad (4.25)$$

*avec  $\tilde{V}^T P \tilde{V}$  minimum*

C'est un système du type précédent (4.20), l'expression de  $\Gamma_{\tilde{V}}$  c'est-à-dire de  $\Gamma_{\tilde{V}}$  s'obtient par (4.23) :

$$\Gamma_{\tilde{V}} = P^{-1} [I - C^T (CP^{-1}C^T)^{-1} CP^{-1}] \sigma_0^2 \quad (4.26)$$

ou encore (4.24) :

$$\Gamma_{\tilde{V}} = {}_n(\delta_{ij})_{2n-r} \begin{pmatrix} P & C^T \\ C & 0 \end{pmatrix}_{(2n-r) \times (2n-r)}^{-1} \cdot {}_{2n-r}(\delta_{ij})_n \sigma_0^2 \quad (4.27)$$

$\delta_{ij} = 1$  si  $i = j \leq n$ ; 0 dans les autres cas

Autrement dit la matrice variance des résidus est la sous-matrice supérieure gauche d'ordre  $n$  de la matrice inverse du système normal.

On peut ensuite si besoin est, calculer  $\Gamma_{\tilde{L}}$  (à partir de la définition  $\tilde{L} = L + \tilde{V}$ ) et  $\tilde{X}$  en exprimant auparavant  $\tilde{X}$  en fonction de  $\tilde{L}$ .

**Application numérique :** Exemple fondamental traité par la méthode des équations de condition. Les calculs ont été faits en 2.5.4.6; du dernier tableau on extrait :

$$\Gamma_{\tilde{V}} = \begin{pmatrix} 2.05 & 2.24 & 0.18 & -0.09 & -0.09 \\ 2.24 & 3.01 & -0.27 & 0.13 & 0.13 \\ 0.18 & -0.27 & 0.41 & -0.20 & -0.20 \\ -0.09 & 0.13 & -0.20 & 0.43 & -0.22 \\ -0.09 & 0.13 & -0.20 & -0.22 & 0.43 \end{pmatrix} \quad \text{avec } \tilde{V} = \begin{pmatrix} \tilde{V}_a \\ \tilde{V}_b \\ \tilde{V}_A \\ \tilde{V}_B \\ \tilde{V}_C \end{pmatrix}$$

Si on calcule alors à partir de  $\Gamma_{\tilde{V}}$  les rapports de corrélation linéaire des variables 2 à 2 ( $r = E[(x_1 - E(x_1)) \cdot (x_2 - E(x_2))] / (\sigma_{x_1} \sigma_{x_2})$ ), on voit par exemple :

- que la liaison entre  $\tilde{V}_a$  et  $\tilde{V}_b$  est presque linéaire ( $r = 2.24 / \sqrt{2.05 \times 3.01} \approx 0.9$ ),
- qu'il n'en est pas du tout de même pour  $a$  et  $B$  ( $r \approx 0.10$ ) : les résidus  $\tilde{V}_a$  et  $\tilde{V}_B$  sont presque non corrélés (indépendants si les observations sont normales).

#### 4.3 PROPRIÉTÉS STATISTIQUES DES ESTIMATEURS DES MOINDRES CARRÉS DANS L'HYPOTHÈSE OÙ LES ERREURS DE MESURES SONT NORMALES ET CENTRÉES

Bien entendu toutes les propriétés et formules établies précédemment dans le cas où on supposait seulement que les erreurs étaient accidentelles restent valables.

L'hypothèse faite à présent est plus forte que la précédente, on suppose à présent :

$$L \in \mathcal{N}(\dot{L}, \sqrt{P^{-1}}\sigma_0)$$

ce qui entraîne évidemment  $E(L - \dot{L}) = 0$ .

Nous allons énoncer en trois théorèmes les propriétés statistiques supplémentaires résultant de la normalité des observations.

**Théorème 4.2** *Les estimateurs fournis par la théorie des moindres carrés sont normaux. Il en est ainsi des estimateurs des inconnues,  $\tilde{X}$ , des résidus vrais  $\tilde{V}$  des observations  $\tilde{L}$ , de toute fonction linéaire des observations vraies.*

On a, en particulier, dans le cas où on utilise des équations d'observations ( $A\dot{X} - H = L + \dot{V}$ ) :

$$\begin{array}{l} \tilde{X} \in \mathcal{N}(\dot{X}, \sqrt{N^{-1}}\sigma_0) \text{ avec } N = A^T P A \\ \tilde{L} \in \mathcal{N}(\dot{L}, \sqrt{AN^{-1}A^T}\sigma_0) \\ \tilde{V} \in \mathcal{N}(0, \sqrt{PUP}\sigma_0) \text{ avec } U = I - \sqrt{P}AN^{-1}A^T\sqrt{P} \end{array} \quad (4.28)$$

(Notons que  $\tilde{L}$  et  $\tilde{V}$  sont des vecteurs normaux dégénérés, respectivement de rang  $r$  et  $(n-r)$ ).

En effet de :

$$\tilde{X} = N^{-1}A^T P(H+L) = N^{-1}A^T PHN^{-1}A^T PL$$

on déduit que  $\tilde{X}$  fonction linéaire du vecteur normal  $L$  est lui-même normal (Cours Statistique, chapitre II); et sa variance a été calculée au paragraphe précédent, c'est  $N^{-1}\sigma_0^2$ .

Quant à  $\tilde{L}$ , on a :  $\tilde{L} = L + \tilde{V} = \dot{L} - \dot{V} + \tilde{V}$ , ce qui, rapproché de (4.8) donne  $\tilde{L} = \dot{L} - AN^{-1}A^T P\dot{V}$  et comme  $\dot{V} = \dot{L} - L \in \mathcal{N}(0, \sqrt{P^{-1}}\sigma_0)$  :

$$\tilde{L} \in \mathcal{N}(\dot{L}, \sqrt{AN^{-1}A^T PP^{-1}\sigma_0^2 PAN^{-1}A^T})$$

soit :  $\tilde{L} \in \mathcal{N}(\dot{L}, \sqrt{AN^{-1}A^T}\sigma_0)$ . Pour  $\tilde{V}$ , et pour toute fonction linéaire de  $\dot{L}$ , ou  $\dot{X}$ , on opérera de façon analogue.

**Théorème 4.3** *Si les observations  $L$  sont normales  $\mathcal{N}(\dot{L}, \sqrt{P^{-1}}\sigma_0)$ , alors la variable  $\chi^2 = (n-r)s_0^2/\sigma_0^2$  où  $s_0^2$  est l'estimation de la variance unitaire  $\sigma_0^2$ , est une variable  $\chi^2$  à  $(n-r)$  degrés de liberté.*

*De plus, si  $\tilde{X}_i$  est la  $i$ ème composante du vecteur  $\tilde{X}$  (estimateur des inconnues) de variance estimée  $s_{\tilde{X}}^2$  alors la variable  $t = (\tilde{X}_i - \dot{X}_i)/s_{\tilde{X}}$  est une variable de Student à  $(n-r)$  degrés de liberté. Plus généralement si  $\tilde{Y}$  est l'estimateur d'une fonction scalaire linéaire des grandeurs observées  $\dot{L}$ , la variable  $t = (\tilde{Y} - \dot{Y})/s_{\tilde{Y}}$  est une variable de Student à  $(n-r)$  degrés de liberté.*

\* - Partons de la formule (4.9) :

$$\sqrt{P}\tilde{V} = U\sqrt{P}\dot{V} \quad (4.29)$$

avec  $U = I - \sqrt{P}AN^{-1}A^T\sqrt{P}$ ,  $U^m = U$ ;  $\text{rang } U = n-r$  et  $(n-r)$  valeurs propres égales à 1 et  $\sqrt{P}\dot{V} \in \mathcal{N}(0, \sigma_0)$ . Il existe un système d'axes, où la représentation de l'opérateur correspondant à la matrice  $U$  est une matrice diagonale  $D$  de la forme :

$${}_n D_n = (\delta_{ij}) \quad \text{avec } \delta_{ij} = \begin{cases} 1 & \text{si } i = j \leq n-r \\ 0 & \text{dans les autres cas} \end{cases}$$

autrement dit, il existe une matrice de rotation  $R$  définissant un changement de variable aléatoire :  $\dot{V} = R(\sqrt{P}\dot{V})$  avec  $\dot{V} \in \mathcal{N}(0, \sigma_0)$  et  $RUR^T = D$  (Cours Statistiques : chapitre I). Il s'ensuit, compte tenu de (4.29) :

$$\tilde{V}^T P \tilde{V} = (\sqrt{P}\dot{V})^T U (\sqrt{P}\dot{V}) = \dot{V} R U R^T \dot{V} = \dot{V}^T D \dot{V}$$

soit :  $\tilde{V}^T P \tilde{V} = \sum_{i=1}^{i=n-r} \tilde{V}_i^2$ .

La quantité  $(n-r)s_0^2 = \tilde{V}^T P \tilde{V}$  est donc la somme de  $(n-r)$  carrés de variables normales indépendantes de même variance  $\sigma_0^2$ .

Il s'ensuit que  $\chi^2 = (n-r)s_0^2/\sigma_0^2$  somme de  $(n-r)$  carrés de variables normales réduites et indépendantes, est une variable  $\chi^2$  à  $(n-r)$  degrés de liberté.

\*- Démontrons à présent la seconde partie du théorème. Soit une quantité scalaire  $\tilde{Y}$  fonction linéaire des observations  $\tilde{L}$  :

$${}_1\tilde{Y}_1 = {}_1D_{mm}\tilde{L}_1 + {}_1F_1$$

L'estimateur de  $\tilde{Y}$ , que fournit la méthode des moindres carrés, est :  $\tilde{Y} = D\tilde{L} + F$ , et sa variance (4.17-4.18) :

$$\sigma_{\tilde{Y}}^2 = D\Gamma_{\tilde{L}}D^T = DAN^{-1}A^TD^T\sigma_0^2$$

On obtient un estimateur sans biais de cette variance,  $s_{\tilde{L}}^2$  en substituant  $s_0^2$  à  $\sigma_0^2$  :

$$s_{\tilde{Y}}^2 = DAN^{-1}A^TD^Ts_0^2 \quad \text{avec } s_0^2 = \frac{1}{n-r}\tilde{V}^TP\tilde{V} \quad (4.30)$$

Considérons alors la variable aléatoire :

$$t = \frac{\tilde{Y} - \hat{Y}}{s_{\tilde{Y}}} = \frac{D(\tilde{L} - \hat{L})}{s_{\tilde{Y}}} \quad (4.31)$$

On voit sans peine qu'on peut écrire :

$$t = \frac{DAN^{-1}A^TP\tilde{V}}{\sigma_{\tilde{Y}}}\frac{\sigma_0}{s_0} = \frac{DAN^{-1}A^TP\tilde{V}}{\sigma_{\tilde{Y}}} \left/ \sqrt{\frac{1}{n-r}\frac{\tilde{V}^T\sqrt{P}U\sqrt{P}\tilde{V}}{\sigma_0^2}} \right. \quad (4.32)$$

En effet  $\tilde{L} - \hat{L} = \tilde{V} - \hat{V}$  et  $\tilde{V} = (I - AN^{-1}A^T)P\tilde{V} \implies \tilde{L} - \hat{L} = AN^{-1}A^TP\tilde{V}$  et d'autre part (4.29) :

$$\sqrt{P}\tilde{V} = U\sqrt{P}\tilde{V} \implies s_0^2 = \frac{1}{n-r}\tilde{V}^T\sqrt{P}U\sqrt{P}\tilde{V} \quad \text{avec } U = I - \sqrt{P}AN^{-1}A^T\sqrt{P}$$

ce qui conduit immédiatement à (4.32). Il est alors facile de montrer que numérateur et dénominateur, variables aléatoires par l'intermédiaire du vecteur aléatoire  $\tilde{V}$ , sont indépendants; il suffit pour cela de démontrer que la variable  $DAN^{-1}A^TP\tilde{V}$  est indépendante du vecteur  $U\sqrt{P}\tilde{V}$  (elle sera alors indépendante de toute fonction de ce dernier vecteur, donc de sa norme), et

pour cela leur covariance est nulle.

C'est en effet pour des vecteurs normaux, ce qui est le cas, une condition nécessaire et suffisante d'indépendance, soit que :

$$E([DAN^{-1}A^T P\dot{V}][(I - \sqrt{P}AN^{-1}A^T \sqrt{P})\sqrt{P}\dot{V}]^T) = 0$$

identité qu'on vérifie immédiatement en tenant compte du fait que  $E(\dot{V}\dot{V}^T) = P^{-1}\sigma_0^2$ .

Il s'ensuit que la variable  $t$  apparaît comme le quotient des 2 variables indépendantes suivantes :

- au numérateur une variable normale réduite :  $(\check{Y} - \check{L}) / \sigma_{\check{Y}}$ ,
- au dénominateur la moyenne quadratique de  $(n - r)$  carrés de variables normales réduites (première partie du théorème),

$t$  est donc une variable de Student à  $(n - r)$  dimensions (degrés de liberté).

**Théorème 4.4** *Si les observations sont normales, les estimateurs des moindres carrés sont totalelement efficaces.*

*Autrement dit, parmi tous les estimateurs sans biais et normaux<sup>1</sup>, ce sont les plus précis (variance minimale).*

Ce théorème capital par ses conséquences théoriques et pratiques découle de la théorie de l'estimation des paramètres d'une loi de probabilité donnée à partir d'un échantillon de variables aléatoires relevant de cette loi.

Il est facile à établir. L'échantillonnage se compose ici des  $n$  composantes du vecteur aléatoire normal :

$${}_n L_1 \in \mathcal{N}(\check{L}, \sqrt{P^{-1}}\sigma_0)$$

La densité de probabilité de la loi est :

---

1. On remarquera qu'en pratique toute méthode de calcul fournit des estimateurs normaux (si les observations sont normales); en effet si  $\check{X}'$  désigne un estimateur de  $\check{X}$ , c'est une fonction des observations  $L$  :  $\check{X}' = f(L) = f(\check{L} + (L - \check{L})) \approx f(\check{L}) +$  fonction linéaire d'erreurs normales; de plus si la méthode de calcul est correcte, c'est-à-dire que par définition elle fournit des estimateurs non biaisés :  $\check{X} = f(\check{L})$ .

$$f(L) = \frac{1}{(2\pi)^{n/2} \cdot (\sigma_0^2)^{n/2}} e^{-\frac{1}{2\sigma_0^2} (L - \hat{L})^T P (L - \hat{L})}$$

soit encore :

$$f(L) = \frac{1}{(2\pi)^{n/2} \cdot (\sigma_0^2)^{n/2}} e^{-\frac{1}{2\sigma_0^2} (A\hat{X} - K)^T P (A\hat{X} - K)}$$

Or d'après la théorie de l'estimation des paramètres inconnus d'une loi donnée (les paramètres inconnus sont ici les  $r$  composantes de  $\hat{X}$ ) à partir d'un échantillon (ici  $L$ ), on sait que la matrice variance de  $\hat{X}$ , estimateur de  $X$ , soit  $\Gamma_{\hat{X}}$  est telle que :

$$Z^T \Gamma_{\hat{X}} Z \geq Z^T F^{-1} Z, \forall Z$$

$F^{-1}$  étant l'inverse de  $F$  matrice d'information de Fisher donnée par :

$$F = -E(F_{ij}) = -E\left(\frac{\partial^2 \text{Log} f}{\partial X_i \partial X_j}\right) \quad (4.33)$$

Or on a ici :  $\text{Log} f = \text{Cte} - \frac{1}{2\sigma_0^2} (A\hat{X} - K)^T P (A\hat{X} - K)$ . D'où on déduit :

$$\frac{\partial \text{Log} f}{\partial X} = \left(\frac{\partial \text{Log} f}{\partial X_i}\right) = -\frac{1}{2\sigma_0^2} (A^T P A \hat{X} - A^T P K) \quad \text{vecteur } r \times 1$$

Puis :

$$\left(\frac{\partial^2 \text{Log} f}{\partial X_i \partial X_j}\right) = -\frac{N}{\sigma_0^2} \implies F = \frac{N}{\sigma_0^2}$$

On a donc :

$$Z^T \Gamma_{\hat{X}} Z \geq Z^T N^{-1} \sigma_0^2 Z, \forall_r Z_1 \quad (4.34)$$

Mais il se trouve que pour l'estimateur des moindres carrés,  $\Gamma_{\hat{X}}$  est justement égal à la matrice variance limite  $N^{-1} \sigma_0^2$ ; on dit alors que l'estimateur des moindres carrés est un estimateur totalemt efficace.

Ce qui signifie que si  $\hat{X}'$  est un autre estimateur non biaisé de  $X$ , on a :

$$\boxed{Z^T \Gamma_{\hat{X}'} Z > Z^T \Gamma_{\hat{X}} Z = Z^T N^{-1} \sigma_0^2 Z, \forall Z}$$



**Conséquences :**

a) Si on considère d'abord une inconnue scalaire  $\dot{Y}$ , alors parmi tous les estimateurs non biaisés et normaux, celui des moindres carrés est de variance minimale.

Autrement dit, si  $\dot{Y}'$  est un autre estimateur non biaisé et normal de  $\dot{Y}$ , on a :

$$\sigma_{\dot{Y}'} > \sigma_{\dot{Y}}$$

ou encore, pour tout  $\varepsilon > 0, \exists N$  tel que :

$$n > N \implies Pr(\|\tilde{Y} - \dot{Y}\| < \varepsilon) > Pr(\|\tilde{Y}' - \dot{Y}\| < \varepsilon)$$

b) Plus généralement si on considère un ensemble de  $r$  grandeurs géométriquement indépendantes  ${}_{,r}\dot{Y}_1$  (les inconnues  ${}_{,r}\dot{X}_1$  sont un cas particulier) dont les estimateurs au sens des moindres carrés sont les composantes du vecteur  ${}_{,r}\dot{Y}_1$  de matrice variance  $\Gamma_{\dot{Y}}$ , alors si  $\tilde{Y}'$  est un autre estimateur non biaisé et normal de  $\dot{Y}$ , on a :

$$Z^T \Gamma_{\dot{Y}} Z > Z^T \Gamma_{\tilde{Y}'} Z, \forall Z \quad (4.35)$$

Si  $Z$  est le vecteur unitaire d'une certaine direction, cette inégalité implique que la variance marginale relative à cette direction est minimale si on choisit l'estimateur des moindres carrés.

Cette inégalité implique que l'ellipsoïde indicateur de la loi relative à l'estimateur des moindres carrés  $\tilde{Y}$  (d'équation  $Z^T \Gamma_{\tilde{Y}}^{-1} Z = 1$ ) est intérieur<sup>a b</sup> à celui relatif à  $\tilde{Y}'$ .

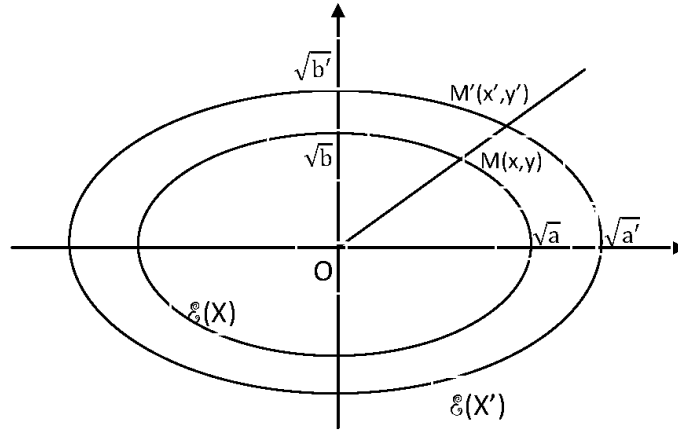
a. Théorème cité par Linnik : je n'en connais pas la démonstration.

b. Yuri Vladimirovich Linnik (1915 - 1972) : mathématicien soviétique était actif en théorie des nombres, probabilité et statistiques mathématiques.

**Théorème 4.2. (Théorème de Linnik[4])** Si deux vecteurs estimateurs normaux et non biaisés  $\tilde{X}$ , et  $\tilde{X}'$  ( $E(\tilde{X}) = E(\tilde{X}')$ ), de même direction et  $\tilde{X}$  est un estimateur des moindres carrés, alors l'ellipsoïde de variance de  $\tilde{X}$  est intérieur à celui relatif à  $\tilde{X}'$ .

Nous présenterons ici la preuve de ce théorème :

*Preuve.* On considère le cas où le vecteur  $\dot{X}$  des inconnues est de dimension 2 et que les inconnues sont indépendantes non corrélées. Soient  $\tilde{X}$  et  $\tilde{X}'$  deux



**Fig. 4.1** Ellipses de variance de  $\tilde{X}$  et  $\tilde{X}'$

estimateurs de  $\dot{X}$  non biaisés ( $E(\tilde{X}) = E(\tilde{X}') = \dot{X}$ ) avec  $\tilde{X}$  est un estimateur des moindres carrés. De 4.35, on écrit :

$$Z^T \Gamma_{\tilde{X}'} Z > Z^T \Gamma_{\tilde{X}} Z, \forall Z \quad (4.36)$$

Pour l'estimateur des moindres carrés  $\tilde{X}$ , on sait que :  $\Gamma_{\tilde{X}} = \sigma_0^2 N^{-1}$ . La matrice  $\sigma_0^2 N^{-1}$  est de la forme :

$$\sigma_0^2 N^{-1} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \quad a > 0, b > 0$$

car  $N$  est symétrique, définie positive. La matrice  $\Gamma_{\tilde{X}'}$  s'écrit :

$$\Gamma_{\tilde{X}'} = \begin{pmatrix} a' & c' \\ c' & b' \end{pmatrix}, \quad a'b' - c'^2 > 0 \implies a'b' > 0$$

car  $\Gamma_{\tilde{X}'}$  est aussi symétrique définie positive. Soit  $Z = (x, y)$ , l'équation (4.36) s'écrit :

$$(x, y) \cdot \begin{pmatrix} a' & c' \\ c' & b' \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} > (x, y) \cdot \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \implies (a' - a)x^2 + 2c'xy + (b' - b)y^2 > 0 \quad (4.37)$$

Soit

$$P(x, y) = (a' - a)x^2 + 2c'xy + (b' - b)y^2 > 0$$

C'est une forme quadratique définie positive, dans ce cas on a nécessairement  $\Delta' = c'^2 - (a' - a)(b' - b) < 0$ . Posons  $\alpha = a' - a, \beta = b' - b$ . D'où  $0 \leq c'^2 < (a' - a)(b' - b) \implies (a' - a)(b' - b) = \alpha.\beta > 0 \implies \alpha > 0$  et  $\beta > 0$  ou  $\alpha < 0$  et  $\beta < 0$ . Or si  $x = 0, y \neq 0 \implies b' - b > 0 \implies b' > b$  ou  $y = 0, x \neq 0 \implies a' - a > 0 \implies a' > a$ . Maintenant on pose  $t = x/y$  avec  $y \neq 0$ , on obtient  $P(x, y) = \frac{1}{y^2}(\alpha t^2 + 2c't + \beta) > 0 \implies \alpha t^2 + 2c't + \beta > 0$ . Comme  $\Delta = c'^2 - \alpha.\beta < 0$ , alors le coefficient  $\alpha$  du polynôme  $\alpha t^2 + 2c't + \beta$  est positif. Il s'ensuit que seul le cas  $a' - a > 0$  et  $b' - b > 0$  est acceptable. Donc  $a' > a$  et  $b' > b$ .

Soit  $\mathcal{E}(\tilde{X})$  l'ellipse de variance de l'estimateur  $\tilde{X}$ . Son équation est donnée par :

$$Z^T \Gamma_{\tilde{X}}^{-1} Z = 1 \quad Z \in \mathbb{R}^2 \quad (4.38)$$

Prenons  $Z = \mathbf{OM} = (x, y)^T, x, y$  étant les composantes du vecteur  $\mathbf{OM}$  (Fig. :4.1), alors (4.38) s'écrit :

$$\frac{x^2}{a} + \frac{y^2}{b} = 1 \quad (4.39)$$

Comme  $a, b$  étant  $>0$  et supposons que  $a > b$ , on a obtenu l'équation d'une ellipse de demi-grand axe  $\sqrt{a}$  et demi-petit axe  $\sqrt{b}$ .

De même, soit  $\mathcal{E}(\tilde{X}')$  l'ellipse de variance de l'estimateur  $\tilde{X}'$ . Son équation est donnée par :

$$Z'^T \Gamma_{\tilde{X}'}^{-1} Z' = 1 \quad Z' \in \mathbb{R}^2 \quad (4.40)$$

Prenons  $Z' = \mathbf{OM}' = (x', y')^T, x', y'$  étant les composantes du vecteur  $\mathbf{OM}'$  (Fig. :4.1), alors (4.40) s'écrit :

$$\frac{x'^2}{a'} + \frac{y'^2}{b'} = 1 \quad (4.41)$$

Comme  $a', b'$  étant  $>0$  et supposons que  $a' > b'$ , on a obtenu l'équation d'une ellipse de demi-grand axe  $\sqrt{a'}$  et demi-petit axe  $\sqrt{b'}$ .

Comme les deux vecteurs estimateurs  $\tilde{X}, \tilde{X}'$  sont colinéaires, alors  $\mathbf{OM}' = \lambda \mathbf{OM}$  avec  $\lambda > 0$ . Comme  $a' > a, b' > b \implies \sqrt{a'} > \sqrt{a}$  et  $\sqrt{b'} > \sqrt{b}$ , alors l'ellipse  $\mathcal{E}(\tilde{X})$  est à l'intérieur de l'ellipse  $\mathcal{E}(\tilde{X}')$ .

Dans le cas où  $\tilde{X}, \tilde{X}'$  sont de dimension supérieure à 2, on obtient que l'ellipsoïde  $\mathcal{E}(\tilde{X})$  est à l'intérieur de l'ellipsoïde  $\mathcal{E}(\tilde{X}')$ .

Fin de la preuve

### **Remarque sur la corrélation des estimateurs**

La méthode des moindres carrés donne les estimateurs les plus précis, mais non nécessairement les moins corrélés.

Considérons par exemple le cas où on cherche à estimer une grandeur inconnue  $\check{X}$  à  $r$  composantes par le système d'équations d'observations suivant :

$$\begin{aligned} X_1 &= l_1 \\ X_2 &= l_2 \\ &\vdots \\ X_r &= l_r \\ X_1 + X_2 + \dots + X_r &= L \end{aligned}$$

Il est évident que si on se limite aux  $r$  premières équations, on obtient un estimateur non biaisé, normal et dont les différentes composantes sont indépendantes et donc non corrélées.

Par contre l'estimateur des moindres carrés, pourtant meilleur que le précédent en ce sens que pour chaque inconnue il donne des écarts-types plus faibles, donne des composantes corrélées 2 à 2.

### **Allure d'un réseau après compensation**

Nous allons voir qu'après compensation tout réseau présente par rapport au réseau vrai une déformation régulière d'allure systématique.

Dans un problème de compensation de réseau, on se propose de déterminer les coordonnées spatiales (parfois seulement les coordonnées planes) de l'ensemble des sommets du réseau  ${}_3\check{X}_{1i} (i = 1, m)$  en compensant toutes les grandeurs observées.

Le réseau est en général "calé" ou "appuyé" sur un certain nombre de points dit d'"appui" en photogrammétrie ou "points anciens" en géodésie.

La question est de savoir comment se situe le réseau compensé  $\check{X}$ , par rapport au réseau vrai  $\check{X}$ , autrement dit qu'elle est l'allure de la "déformée" du réseau compensé par rapport au réseau vrai.

Considérons en particulier les écarts relatifs aux côtés :  $\check{z}_i - \dot{z}_i$  (le raisonnement est valable pour les autres coordonnées).

On s'attendrait assez naturellement à ce que ces écarts se répartissent aléatoirement autour de 0, c'est-à-dire à ce que la déformée présente l'allure suivante :

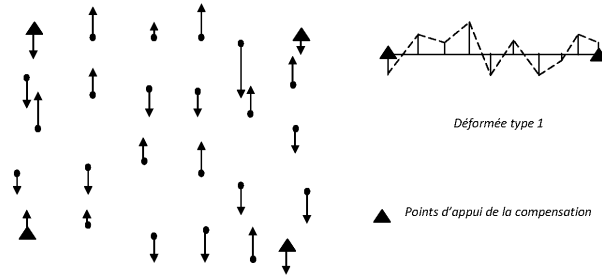


Fig. 4.2 La déformée type 1

En fait il n'en est rien, et assez paradoxalement à prime abord, les écarts présentent un systématisme : on obtiendra par exemple l'allure suivante pour la déformée :

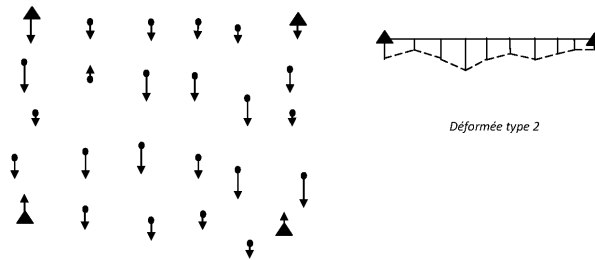


Fig. 4.3 La déformée type 2

Il est extrêmement important de comprendre la raison de ceci, et le fait que contrairement aux apparences, un tel résultat est tout à fait logique, et meilleur que le précédent :

- il n'y a d'abord aucune raison pour que les écarts se répartissent aléatoirement autour de la valeur 0 : les inconnues en effet sont corrélées, et d'autant plus en général qu'elles sont relatives à des sommets plus proches dans l'espace,
- mais il y a plus : les estimateurs au sens des moindres carrés sont à variance minimale; il en est donc ainsi pour la courbure locale en chaque point de la déformée  $\tilde{X} - \hat{X}$ , qui est l'un d'eux et dont la vraie valeur est nulle.

Or il est bien clair que la variance des courbures locales des déformées aléatoires du type 2 sont bien plus faibles que les variances des courbures locales des déformées aléatoires du type 1.

Il est donc tout à fait logique que la solution au sens des moindres carrés présente par rapport à la solution vraie une déformation d'allure systématique : il s'agit en fait d'un pseudo-systématisme qui n'a rien à voir avec le systématisme induit par des erreurs systématiques dont il convient de le distinguer soigneusement, et qui traduit une supériorité de la méthode des moindres carrés sur les autres méthodes d'estimation : le réseau compensé par moindres carrés est celui qui statistiquement présente le moins de discontinuités d'écart par rapport au réseau vrai.

#### 4.4 COMPARAISON DE LA VARIANCE UNITAIRE A PRIORI $\sigma_0^2$ QUAND ELLE EST CONNUE AVEC LA VARIANCE UNITAIRE ESTIMÉE $s_0^2$

Le test permet de voir si l'hypothèse faite au départ, à savoir que le vecteur  ${}_nL_1$  des observations obéit à la loi normale  $\mathcal{N}(\bar{L}, \sqrt{P^{-1}}\sigma_0)$  est à rejeter ou non.

Si cette hypothèse est à rejeter, cela peut provenir de diverses raisons :

- mauvais choix de  $\sigma_0$  ou des poids des diverses catégories d'observations,
- les observations ne sont pas normales.

Le test est basé sur le théorème 4.3 selon lequel, si les observations sont normales, la variable aléatoire :

$$\chi^2 = (n-r) \frac{s_0^2}{\sigma_0^2} \quad (4.42)$$

est une variable  $\chi^2$  à  $(n-r)$  degrés de liberté ( $n$  nombre de grandeurs observées,  $r$  nombre d'inconnues).

On se donne alors un seuil  $\eta$  des probabilités négligeables, et on détermine avec la table du  $\chi^2$  la valeur  $\chi_\eta^2$  telle que :

$$Pr(\chi^2 > \chi_\eta^2) = \eta$$

Si le  $\chi^2$  calculé (4.42) est supérieur à  $\chi_\eta^2$ , on rejette alors l'hypothèse que les observations  $L$  sont normales  $\mathcal{N}(L, \sqrt{P^{-1}}\sigma_0)$ .

**Exemple :**

Dans l'exemple fondamental 1.2, où il y a deux catégories d'observations, on a pris  $\sigma_0^2 = 1$ . Pour estimation de  $\sigma_0^2$  on a trouvé :  $s_0^2 = 0.68$ . On calcule :

$$\chi^2 = (n - r) \frac{s_0^2}{\sigma_0^2} = (5 - 3) \times \frac{0.68}{1} = 1.36$$

Pour un seuil  $\eta = 5\%$ , la table  $\chi^2$  à 2 degrés de liberté donne :

$$\chi_\eta^2 = 6$$

Comme  $\chi^2 < \chi_\eta^2$ , l'hypothèse faite (choix de la loi des erreurs d'observations) est acceptable, sous réserve que d'autres tests ne l'infirmant pas.





# Chapitre 5

## Précision, Justesse et Exactitude d'un Ensemble de Grandeurs Estimées par la Méthode des Moindres Carrés

### Sommaire

|       |  |     |
|-------|--|-----|
| 5.1   | INTRODUCTION.....  | 96  |
| 5.1.1 | Définition de l'erreur .....   | 97  |
| 5.2   | ESTIMATIONS PONCTUELLES DE LA PRÉCISION, DE LA JUSTESSE ET DE L'EXACTITUDE DE L'ESTIMATEUR D'UNE GRANDEUR A $r$ DIMENSIONS DANS LE CADRE D'UN SYSTÈME $S$ DE MESURES.....    | 98  |
| 5.2.1 | Précision de l'estimateur $\hat{X}$ .....  | 98  |
| 5.2.2 | Justesse de l'estimateur $\hat{X}$ .....   | 100 |
| 5.2.3 | Exactitude de l'estimateur $\hat{X}$ .....   | 100 |
| 5.3   | ESTIMATION PONCTUELLE DE LA PRÉCISION DE LA VARIANCE ESTIMÉE ET DE L'ECART-TYPE ESTIMÉ DE LA MESURE DE POIDS UNITAIRE (DANS LE CAS OÙ LA REDONDANCE $n - r$ EST FORTE) ..... | 102 |
| 5.4   | ESTIMATION DE LA PRÉCISION ET DE L'EXACTITUDE DES ESTIMATEURS PAR LA MÉTHODE DES INTERVALLES DE CONFIANCE - INTERVALLE DE CONFIANCE POUR L'ECART-TYPE UNITAIRE .....         | 103 |

|       |   |     |
|-------|---|-----|
| 5.4.1 | Intervalle de confiance pour l'espérance de l'estimateur d'une grandeur inconnue scalaire dans le cadre d'un certain système $S$ de mesures . . . . . | 104 |
| 5.4.2 | Intervalle de confiance pour la valeur universelle d'une grandeur inconnue scalaire . . . . .   | 104 |
| 5.4.3 | Intervalle de confiance pour l'écart-type unitaire $\sigma_0$ .   | 105 |

---

En ce qui concerne les estimations ponctuelles de la précision, de la justesse et de l'exactitude, la question a déjà été traitée en détail dans le chapitre sur les erreurs ; nous rappellerons ici l'essentiel pour la commodité.

On appliquera ensuite aux mêmes problèmes la méthode des intervalles de confiance.

## 5.1 INTRODUCTION

Jusqu'ici, dans ce chapitre nous avons fait des raisonnements ne découlant que des hypothèses suivantes :

- à chaque grandeur scalaire observée ou inconnue on peut faire correspondre un nombre, qui est "sa vraie valeur".
- toutes les observations  ${}_nL_1$  sont non biaisées :

$$E({}_nL_1) = {}_n\dot{L}_1 = \text{vraies valeurs des grandeurs observées}$$

autrement dit les erreurs commises lors des mesures sont des erreurs d'observations,

- on connaît à un facteur scalaire près la matrice variance des observations  $\Gamma_L = P^{-1}\sigma_0^2$ ,
- on peut sans inconvénient linéariser le problème (on connaît donc des valeurs approchées suffisamment bonnes pour chaque grandeur inconnue) : après linéarisation, il existe alors une relation matricielle linéaire entre les variations  $\dot{X}$  des vraies valeurs des  $r$  inconnues, et les vraies valeurs  $\dot{L}$  des  $n$  grandeurs observées :

$$A\dot{X} - (H + \dot{L}) = 0$$

(Nous nous limiterons ici à la considération de la seule méthode des équations d'observations, le lecteur généralisera sans peine aux autres cas).

On a alors obtenu pour les inconnues  $\hat{X}$  en minimisant la somme des carrés pondérés des résidus, l'estimateur :

$$\hat{X} = N^{-1}A^T PL \quad \text{avec} \quad N = A^T PA$$

qui est sans biais ( $E(\hat{X}) = \hat{X}$ ) et qui a pour matrice variance :  $\Gamma_{\hat{X}} = N^{-1}\sigma_0^2$ .

Mais nous avons vu (chapitre 1) que ces hypothèses acceptées telles quelles conduisent à croire que tout accroissement du nombre de mesures se traduit par une amélioration de la concordance entre les estimations et les "vraies valeurs" des grandeurs estimées. Or ceci est infirmé par l'expérience.

On peut comme nous l'avons vu en donner diverses raisons : en particulier, on ne peut définir la "vraie valeur d'une grandeur" et d'autre part, même en l'absence de systématisme, les instruments en dépit du nombre d'observations qu'on peut faire sont incapables de distinguer des grandeurs trop voisines (pouvoir séparateur).

### 5.1.1 Définition de l'erreur

Nous avons alors été amenés, dans le cas de l'estimation d'une grandeur à une dimension selon un certain type de processus de mesure, à substituer à la notion de valeur vraie de la grandeur, celle de valeur universelle relative à ce type de processus de mesures, notée  $\dot{L}$ , moyenne d'un très grand nombre de mesures effectuées avec un grand nombre de processus optimaux de ce type; on peut alors définir l'erreur  $e$  comme la différence entre l'estimation  $L$  et la valeur universelle  $\dot{L}$  :

$$e = L - \dot{L}$$

Dans le cadre d'un certain processus de mesures,  $L$  a une espérance  $E(L)$  distincte en général de  $\dot{L}$ , et on peut écrire :

$$e = (L - E(L)) + (E(L) - \dot{L}) = \varepsilon + \beta$$

avec  $\varepsilon = L - E(L)$  et  $E(\varepsilon) = 0$ ,  $\beta = E(L) - \dot{L}$

**Définition 5.1 (Définition de l'erreur)** *L'erreur est somme d'une erreur d'observation résiduelle  $\varepsilon$  et d'un biais  $\beta$ . La précision a alors été chiffrée par l'écart-type de l'erreur d'observation résiduelle, la justesse d'un tel système par le biais moyen quadratique*

$\sqrt{E(\beta^2)}$  et l'exactitude par l'erreur moyenne quadratique (somme quadratique de la précision et de la justesse).

Nous avons vu enfin, que pour des méthodes et appareillages bien adaptés à leur but, il y avait de bonnes raisons de penser que le biais moyen quadratique était voisin de l'écart-type de l'observation élémentaire!

Ces définitions et propositions peuvent être généralisées dans le cas de l'estimation d'une grandeur à  $r$  dimensions ( $r$  grandeurs inconnues) au moyen d'une grandeur observée à  $n$  dimensions.

## 5.2 ESTIMATIONS PONCTUELLES DE LA PRÉCISION, DE LA JUSTESSE ET DE L'EXACTITUDE DE L'ESTIMATEUR D'UNE GRANDEUR A $r$ DIMENSIONS DANS LE CADRE D'UN SYSTÈME S DE MESURES

### 5.2.1 Précision de l'estimateur $\tilde{X}$

**Définition 5.2** La précision de  $\tilde{X}$  juge l'accord entre  $\tilde{X}$  et  $E(\tilde{X})$  (espérance de  $\tilde{X}$  dans le cadre du système  $S$ ); elle est décrite par la matrice variance  $\Gamma_{\tilde{X}}$  de  $\tilde{X}$  qui caractérise la fluctuation de  $\tilde{X}$  autour de  $E(\tilde{X})$ .

Rappelons que dans le cas où on utilise la méthode des équations d'observations, on a :

$$\Gamma_{\tilde{X}} = N^{-1} \sigma_0^2$$

où  $\sigma_0^2$  est la variance unitaire; dans le cas où la variance unitaire est inconnue, on lui substitue son estimateur  $s_0^2 = \frac{V^T P V}{(n-r)}$  et on obtient pour  $\Gamma_{\tilde{X}}$  l'estimateur sans biais :

$$\tilde{\Gamma}_{\tilde{X}} = N^{-1} s_0^2$$

Rappelons aussi certains usages qui peuvent être faits de la connaissance de  $\Gamma_{\tilde{X}}$  si les observations  $L$  peuvent être considérées comme normales; dans ce cas, la loi du vecteur aléatoire  $\tilde{X}$  est normale :

$$\tilde{X} \in \mathcal{N}(E(\tilde{X}), \sqrt{\Gamma_{\tilde{X}}})$$

autrement dit :

$$Pr(x < \tilde{X} < x + dx) = \frac{1}{(\sqrt{2\pi})^n \sqrt{\Gamma}} e^{-\frac{1}{2}(x - E(\tilde{X}))^T \Gamma_{\tilde{X}}^{-1} (x - E(\tilde{X}))} dx$$

L'ellipsoïde indicateur de la loi a pour équation :

$$(x - E(\tilde{X}))^T \Gamma_{\tilde{X}}^{-1} (x - E(\tilde{X})) = 1$$

et la probabilité pour que  $\tilde{X}$  soit intérieur à l'ellipsoïde  $E_{\rho^2}$  d'égale densité d'équation :

$$(x - E(\tilde{X}))^T \Gamma_{\tilde{X}}^{-1} (x - E(\tilde{X})) = \rho^2$$

est donnée par la loi du  $\chi^2$  à  $(n - r)$  degrés de liberté :

$$Pr(\tilde{X} \in E_{\rho^2}) = Pr(\chi^2 < \rho^2)$$

**Exemple :**

Considérons le cas de l'exemple fondamental; la précision du vecteur estimateur des inconnues  $a, b, c$  (côtés du triangle) est décrite par sa matrice variance.

L'estimation  $\tilde{\mathcal{X}}$  de  $\mathcal{X}$  est :

$$\tilde{\mathcal{X}} = \mathcal{X}_0 + \tilde{X}$$

( $\mathcal{X}_0$  valeur approchée;  $\tilde{X}$  estimation fournie par les moindres carrés). D'où :

$$\Gamma_{\tilde{\mathcal{X}}} = \Gamma_{\tilde{X}} = N^{-1} \sigma_0^2 = N^{-1} \text{ (si } \sigma_0^2 = 1 \text{)}$$

Soit (2.5.4.6) :

$$\Gamma_{\tilde{\mathcal{X}}} = \begin{pmatrix} 2.06 & 2.24 & 1.08 \\ 2.24 & 3.00 & 1.85 \\ 1.08 & 1.85 & 2.12 \end{pmatrix} \quad (5.1)$$

En particulier on obtient comme estimations des précisions pour chacun des 3 côtés les écarts-types suivants exprimés en *dcm*m (l'unité utilisée dans les calculs est 0.1 mm pour les longueurs) :

$$\begin{aligned} \sigma_{\tilde{a}} &= \sqrt{2.06} \approx 1.44 \text{ dcm}m \\ \sigma_{\tilde{b}} &= \sqrt{3.00} \approx 1.73 \text{ dcm}m \\ \sigma_{\tilde{c}} &= \sqrt{2.12} \approx 1.46 \text{ dcm}m \end{aligned}$$

Chiffres inférieurs quant aux côtés mesurés  $a$  et  $b$  aux *emq* correspondantes ( $emq_a = 1.9 \text{ dcm}m$  et  $emq_b = 2.5 \text{ dcm}m$ ), mais pas tellement : le rapport du

nombre de grandeurs observées au nombre d'inconnues n'est pas en effet très grand.

Proposons nous à présent d'estimer la précision des observations compensées  $\tilde{L}$ ; on a (4.18) :

$$\Gamma_{\tilde{L}} = AN^{-1}A^T \sigma_0^2 = AN^{-1}A^T$$

soit tous calculs faits :

$$\Gamma_{\tilde{L}} = \begin{pmatrix} 2.06 & 2.24 & 0.19 & -0.10 & -0.09 \\ 2.24 & 3.01 & -0.28 & 0.14 & 0.14 \\ 0.19 & -0.28 & 0.42 & -0.21 & -0.21 \\ -0.10 & 0.14 & -0.21 & 0.42 & -0.23 \\ -0.09 & 0.14 & -0.21 & -0.23 & 0.44 \end{pmatrix}$$

En particulier on trouve comme précision des angles compensés :

$$\sigma_{\tilde{A}_i} = 0.65 \text{ dcmgr} \text{ (pour } emq_{A_i} = 0.81 \text{ dcmgr)} \quad A_i = \hat{A}, \hat{B}, \hat{C}$$

### 5.2.2 Justesse de l'estimateur $\tilde{X}$

**Définition 5.3** La justesse de  $\tilde{X}$  est l'accord entre  $E(\tilde{X})$  et  $\tilde{X}$ , c'est-à-dire entre l'espérance de  $\tilde{X}$  dans le cadre du système  $S$  de mesures, et la valeur universelle de  $\tilde{X}$  de la grandeur inconnue; on le chiffre par le biais  $\beta$  différence entre  $E(\tilde{X})$  et  $\tilde{X}$ .

Statistiquement nous le caractériserons par  $E(\beta\beta^T)$  matrice variance du biais pour tous les systèmes de caractéristiques voisines. En l'absence d'indication sur le biais, il y a de bonnes raisons d'adopter pour décrire la justesse la matrice variance des déterminations élémentaires :

$$\Gamma_{DE} \approx \frac{n}{r} \Gamma_{\tilde{X}}$$

( $n$  nombre des grandeurs observées;  $r$  nombre d'inconnues).

Rappelons qu'une détermination élémentaire de  $\tilde{X}$ , est l'une des déterminations qu'on peut calculer à partir de  $r$  grandeurs observées seulement; en supposant également probables tous les types de déterminations élémentaires qu'on peut constituer à partir des  $n$  grandeurs observées, on définit

une loi de probabilité dont la matrice variance est  $\approx \Gamma_{DE}$ .  $\Gamma_{DE}$  joue pour une grandeur à  $r$  dimensions, le rôle que joue la variance d'une mesure élémentaire pour une grandeur à une dimension. (Voir pour détails et raisonnements le chapitre sur les erreurs).

### 5.2.3 Exactitude de l'estimateur $\tilde{X}$

Par définition l'erreur vaut :

$$\begin{aligned}
 e &= \tilde{X} - \dot{X} = (\tilde{X} - E(\tilde{X})) + (E(\tilde{X}) - \dot{X}) = \varepsilon + \beta \\
 \text{avec : } \quad \varepsilon &= \text{erreur d'observation résiduelle} \\
 \beta &= \text{biais du système } S \text{ de mesures}
 \end{aligned}
 \tag{5.2}$$

Si nous considérons à présent tous les systèmes de mesures de même précision  $\Gamma$  permettent d'estimer  $\dot{X}$ , on peut admettre que le biais  $\beta$  est lui-même une variable aléatoire centrée, indépendante de  $\varepsilon$ .

Par suite :  $E(ee^T) = E(\varepsilon\varepsilon^T) + E(\beta\beta^T)$ . Soit :

$$E(ee^T) = \Gamma + E(\beta\beta^T)$$

C'est par définition la matrice-variance des erreurs relatives à  $\tilde{X}$ . C'est avec elle que nous chiffrerons l'exactitude de l'estimateur.

**Définition 5.4** *L'exactitude d'une estimation obtenue dans le cadre d'un système de mesures  $S$  indique l'accord entre cette estimation et la valeur universelle  $\dot{X}$  de la grandeur estimée. Statistiquement on le chiffrera par la matrice-variance des erreurs,  $E(ee^T)$  somme de la matrice variance des erreurs d'observations résiduelles et de la matrice variance du biais; cette matrice indique comment fluctue l'estimateur  $\tilde{X}$  autour de  $\dot{X}$ .*

Dans le cas de systèmes  $S$  de mesures bien élaborées (systématismes négligeables), nous avons vu qu'on peut prendre pour caractériser la justesse la matrice variance des déterminations élémentaires,  $\Gamma_{DE}$  et à défaut de la connaissance de celle-ci, la matrice  $\frac{n}{r} \cdot \Gamma_{\bar{X}}$  :

$$E(ee^T) = \left(1 + \frac{n}{r}\right) \Gamma_{\bar{X}} \approx \left(1 + \frac{n}{r}\right) \Gamma_{DE}$$

On voit que l'exactitude maximale est limitée par le biais; on a au mieux en moyenne :

$$E(ee^T) = E(\beta\beta^T)$$

et pour un système de mesures bien élaborée au mieux (en moyenne) :

$$E(ee^T) = \frac{n}{r} \Gamma_{\bar{X}} \approx \Gamma_{DE}$$

**Exemple :**

Appliquons ce qui précède au cas de l'exemple fondamental; on a ici (3 inconnues, 5 grandeurs observées), tenant compte de 2.2.1.1 :

$$E(ee^T) = \left(1 + \frac{5}{3}\right) \cdot \Gamma_{\bar{X}} = \begin{pmatrix} 5.49 & 5.97 & 2.88 \\ 5.97 & 8.00 & 4.93 \\ 2.88 & 4.93 & 5.65 \end{pmatrix}$$

D'où, en nous rappelant que les calculs ont été faits en prenant comme unité de longueurs le  $dcmm$  :

$$emq_a = \sqrt{5.49} dcmm = 0.23 mm$$

$$emq_b = \sqrt{8.00} dcmm = 0.28 mm$$

$$emq_c = \sqrt{5.65} dcmm = 0.24 mm$$

On observera que les exactitudes ainsi estimées pour les côtés compensés  $\tilde{a}$  et  $\tilde{b}$  sont très voisines de celles estimées pour les mesures  $a$  et  $b$  ( $emq_a = 0.19 mm$  et  $emq_b = 0.25 mm$ ).

D'une façon analogue, on obtiendra en multipliant par le facteur  $(1 + 5/3)$  la variance correspondante, une estimation de la variance d'erreur de toute fonction linéaire des grandeurs observées.

Ainsi on a trouvé pour les angles, une précision :  $\sigma_{\bar{\alpha}_i} = 0.65 dcmmgr$ .



L'exactitude correspondante sera chiffrée par :  $emq_{A_i} = \sqrt{1 + 5/3} \times 0.65 = 1.06 dcmgr$ .

### 5.3 ESTIMATION PONCTUELLE DE LA PRÉCISION DE LA VARIANCE ESTIMÉE ET DE L'ÉCART-TYPE ESTIMÉ DE LA MESURE DE POIDS UNITAIRE (DANS LE CAS OÙ LA REDONDANCE $n - r$ EST FORTE)

On remarquera qu'il n'y a pas de parler de l'exactitude d'un écart-type : un écart-type n'est pas une grandeur physique mais une grandeur mathématique statistique, caractéristique de la loi de probabilité du système  $S$  de mesures et dont on peut obtenir pratiquement une estimation aussi parfaite que l'on veut simplement en multipliant suffisamment le nombre des mesures.

Pour évaluer la précision de la variance et l'écart-type unitaires estimés  $s_0^2$  et  $s_0$  nous partirons du fait que la variable :

$$\chi^2 = (n - r) \frac{s_0^2}{\sigma_0^2}$$

( $\sigma_0^2$  variance unitaire) suit la loi  $\chi^2$  à  $(n - r)$  degrés de liberté (théorème 4.3 du paragraphe 4.3).

Il s'ensuit, que si  $(n - r)$  est suffisamment grand  $\chi^2$  suit la loi limite du  $\chi^2$  qui est comme on sait une loi normale de moyenne  $(n - r)$  et d'écart-type  $\sqrt{2(n - r)}$  :

$$(n - r)s_0^2 / \sigma_0^2 \in \mathcal{N} \left( n - r, \sqrt{2(n - r)} \right)$$

d'où on déduit :

$$(n - r)(s_0^2 / \sigma_0^2 - 1) \in \mathcal{N} \left( 0, \sqrt{2(n - r)} \right)$$

$$s_0^2 \in \mathcal{N} \left( \sigma_0^2, \sqrt{\frac{2}{n - r}} \sigma_0^2 \right)$$

et par suite l'écart-type de l'estimateur de la variance unitaire  $s_0^2$ , peut être estimé par :

$$\sigma_{s_0}^2 = \sqrt{\frac{2}{n-r}} \sigma_0^2 \quad (5.3)$$

Dans la pratique, on recourt plutôt pour chiffrer la précision d'une précision, à l'écart-type estimé de l'écart-type estimé, soit  $\sigma_{s_0}$ . Pour l'évaluer nous poserons :  $s_0 = \sigma_0 + \varepsilon$ .

Soit, en remarquant que  $\varepsilon$  est petit devant  $\sigma_0$ , puisque la redondance  $(n-r)$  des observations est supposée forte :

$$\begin{aligned} s_0^2 &\approx \sigma_0^2 + 2\sigma_0\varepsilon \\ \varepsilon &\approx \frac{s_0^2 - \sigma_0^2}{2\sigma_0} \end{aligned}$$

D'où il suit, en désignant par  $\sigma_{s_0}^2$  la variance de  $s_0$  :

$$\sigma_{s_0}^2 \approx E(s_0 - \sigma_0)^2 = E(\varepsilon^2) \approx \frac{E(s_0^2 - \sigma_0^2)^2}{4\sigma_0^2}$$

Puis, tenant compte du fait que  $E(s_0^2 - \sigma_0^2)^2$  est la variance de  $s_0^2$ ,  $\sigma_0^2$  estimée en (5.3) :

$$\sigma_{s_0}^2 \approx \frac{(2/(n-r))\sigma_0^4}{4\sigma_0^2} = \frac{\sigma_0^2}{2(n-r)}$$

D'où on déduit enfin :

$$\sigma_{s_0} = \frac{s_0}{\sqrt{2(n-r)}} \quad (5.4)$$

#### 5.4 ESTIMATION DE LA PRÉCISION ET DE L'EXACTITUDE DES ESTIMATEURS PAR LA MÉTHODE DES INTERVALLES DE CONFIANCE - INTERVALLE DE CONFIANCE POUR L'ÉCART-TYPE UNITAIRE

La méthode des intervalles de confiance est surtout intéressante dans le cas où d'une part le degré de surabondance du nombre  $n$  de grandeurs mesurées par rapport au nombre  $r$  de grandeurs inconnues est faible, c'est-à-dire si la redondance  $(n-r)$  est faible et où d'autre part la variance unitaire n'est pas connue a priori.

Dans les autres cas la variance unitaire est connue avec suffisamment de précision pour que les estimations ponctuelles données au paragraphe précédent soit suffisamment sûres.

Nous supposerons dans tout ce qui suit, les observations normales.

#### **5.4.1 Intervalle de confiance pour l'espérance de l'estimateur d'une grandeur inconnue scalaire dans le cadre d'un certain système $S$ de mesures**

Soit  $\tilde{X}_i$  l'estimateur de la valeur de la grandeur et  $E(\tilde{X}_i)$  l'espérance mathématique de  $\tilde{X}_i$  dans le cadre du système de mesures  $S$ ; soit  $s_{\tilde{X}_i}$  l'écart-type estimé de  $\tilde{X}_i$ .

Si  $n$  est le nombre de grandeurs observées et  $r$  le nombre de grandeurs inconnues, on sait d'après le théorème 4.3 du paragraphe 4.3 que la variable :

$$t = \frac{\tilde{X}_i - E(\tilde{X}_i)}{s_{\tilde{X}_i}}$$

est une variable de Student à  $(n - r)$  degrés de liberté.

Pour déterminer un intervalle de confiance à  $(1 - \eta)\%$  on cherche alors dans la table de la loi de Student et pour  $(n - r)$  degrés la valeur  $t_\eta$  telle que :

$$Pr(|t| > t_\eta) = \eta$$

Il y a alors  $(1 - \eta)\%$  de chances pour que :

$$\tilde{X}_i - t_\eta s_{\tilde{X}_i} < E(\tilde{X}_i) < \tilde{X}_i + t_\eta s_{\tilde{X}_i}$$

#### **5.4.2 Intervalle de confiance pour la valeur universelle d'une grandeur inconnue scalaire**

Nous n'envisagerons ici que le cas de systèmes de mesures où les systématismes sont bien maîtrisés; dans ce cas nous avons vu qu'on pouvait estimer la matrice variance d'erreur de l'estimateur d'une grandeur à  $r$  dimen-

sions en multipliant la matrice de variance des moindres carrés par le facteur  $\left(1 + \frac{n}{r}\right)$ .

Il s'ensuit que si  $\sigma_{\tilde{X}_i}$  est l'écart-type de la grandeur scalaire de valeur universelle  $\dot{X}_i$ , l'emq de l'estimateur  $\tilde{X}_i$  de  $\dot{X}_i$  est :

$$emq_{\tilde{X}_i} = \sqrt{1 + \frac{n}{r}} \cdot \sigma_{\tilde{X}_i}$$

Par suite :  $\tilde{X}_i \in \mathcal{N}(\dot{X}_i, \sqrt{1 + \frac{n}{r}} \cdot \sigma_{\tilde{X}_i})$ .

Comme on ne connaît pas  $\sigma_{\tilde{X}_i}$  en général, on lui substitue l'estimation :  $s_{\tilde{X}_i} = (\sigma_{\tilde{X}_i} / \sigma_0) s_0$ . Dans ces conditions, la variable :

$$t = \frac{\tilde{X}_i - \dot{X}_i}{\sqrt{1 + \frac{n}{r}} \cdot s_{\tilde{X}_i}} = \frac{\tilde{X}_i - \dot{X}_i}{\sqrt{1 + \frac{n}{r}} \cdot \sigma_{\tilde{X}_i}} \Big/ \frac{s_0}{\sigma_0}$$

apparaît comme une variable de Student à  $(n - r)$  degrés de liberté : les raisonnements à faire sont à peu de choses près ceux qui ont été faits au paragraphe 4.3.

L'intervalle de confiance de  $\dot{X}$  à  $(1 - \eta)\%$  est alors :

$$\left[ \tilde{X}_i - \sqrt{1 + \frac{n}{r}} \cdot t_\eta s_{\tilde{X}_i}, \tilde{X}_i + \sqrt{1 + \frac{n}{r}} \cdot t_\eta s_{\tilde{X}_i} \right]$$

$t_\eta$  étant la valeur de  $t$ , telle que :  $Pr(|t| > t_\eta) = \eta$ .

### 5.4.3 Intervalle de confiance pour l'écart-type unitaire $\sigma_0$

On s'appuie sur le fait que la variable :

$$\chi^2 = (n - r) s_0^2 / \sigma_0^2$$

suit la loi du  $\chi^2$  à  $(n - r)$  degrés de liberté; le raisonnement à faire est un raisonnement classique de statistique et ne sera pas refait ici.

La table VIII (cours de statistique) donne pour  $\eta$  donné et pour  $(n - r)$  degrés de liberté les nombres  $\gamma_1$  et  $\gamma_2$  tels qu'il y ait  $(1 - \eta)\%$  de chances pour que :

$$\gamma_1 s_0 < \sigma_0 < \gamma_2 s_0$$

# Annexe A : Rappels d'Éléments de Calcul Matriciel

## 5.5 LES APPLICATIONS LINÉAIRES

**Définition 5.5** Soient  $U, V$  deux espaces vectoriels sur  $\mathbb{R}$ . Une application linéaire  $f$  de  $U \Rightarrow V$  est dite linéaire si et seulement si :

$$\forall \mathbf{x}, \mathbf{y} \in U, \quad f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}) \quad (5.5)$$

$$\forall \mu \in \mathbb{R}, \mathbf{x} \in U \quad f(\mu \cdot \mathbf{x}) = \mu f(\mathbf{x}) \quad (5.6)$$

si  $U = V$ ,  $f$  est un endomorphisme de  $U$ .

L'application  $f$  est dite :

- surjective si  $\forall \mathbf{y} \in V, \exists \mathbf{x} \in U$  et  $\mathbf{y} = f(\mathbf{x})$ ,
- injective si  $\mathbf{x} \neq \mathbf{y}$  alors  $f(\mathbf{x}) \neq f(\mathbf{y})$ ,
- bijective si  $f$  est surjective et injective ou encore que  $\mathbf{y} = f(\mathbf{x})$  a une solution.

$\text{Ker} f = \{\mathbf{x} \in U / f(\mathbf{x}) = \mathbf{0} =\} =$  le noyau de  $f$ .

$$\text{Im}f = \{\mathbf{y} \in V / \exists \mathbf{x} \in U \text{ et } f(\mathbf{x}) = \mathbf{y}\}.$$

Si  $V = U$  et  $f$  bijective, alors  $f$  est un automorphisme de  $U$ .

On considère  $U$  et  $V$  deux espaces vectoriels de dimension finie c'est-à-dire que les bases de  $U$  et de  $V$  sont finies.

Soit  $f$  une application linéaire de  $\mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $f$  est définie par la donnée de  $f(\mathbf{e}_i)$  pour  $i = 1, 2, \dots, n$  et  $(\mathbf{e}'_j)$   $j = 1, 2, \dots, m$  la base de  $\mathbb{R}^m$ , alors :

$$f(\mathbf{e}_i) = a_{1i} \cdot \mathbf{e}'_1 + a_{2i} \cdot \mathbf{e}'_2 + \dots + a_{mi} \cdot \mathbf{e}'_m \quad (5.7)$$

On a donc le tableau suivant :

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2i} & \cdots & a_{2n} \\ \vdots & \cdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mi} & \cdots & a_{mn} \end{pmatrix} \quad (5.8)$$

Le tableau  $A = (a_{ij})$  s'appelle matrice à  $m$  lignes et  $n$  colonnes.

Si  $m = n$ ,  $A$  est dite matrice carrée d'ordre  $n$ . Dans la suite, on considère les matrices carrées d'ordre  $n$ .

## 5.6 OPÉRATIONS SUR LES MATRICES

- Soit  $A = (a_{ij})$ , les éléments  $a_{ii}$  sont les éléments diagonaux. Les éléments  $a_{ij}$  avec  $i \neq j$  sont les éléments non-diagonaux.

- Soit  $A = (a_{ij})$  et  $B = (b_{ij})$  et  $C = A + B$ . L'élément  $c_{ij}$  de  $C$  est tel que :

$$c_{ij} = a_{ij} + b_{ij} \quad i, j = 1, 2, \dots, n$$

- Soit  $C = \mu A$  où  $\mu$  est un réel, alors  $C = (c_{ij})$  avec  $c_{ij} = \mu \cdot a_{ij}$  où  $i, j = 1, 2, \dots, n$ .

- Soit  $C = A \cdot B$  le produit de 2 matrices carrées, alors  $C = (c_{ij})$  telle que :

$$c_{ij} = \sum_{k=1}^{k=n} a_{ik} b_{kj} \quad (5.9)$$

on a en général :

$$A.B \neq B.A$$

- La matrice  $O = (0)$  matrice dont tous les éléments sont nuls est la matrice neutre pour l'addition :

$$A + O = O + A = A$$

- La matrice unité  $I = (\delta_{ij})$  avec  $\delta_{ij} = 1$  si  $i=j$  et  $\delta_{ij} = 0$  si  $i \neq j$  c-à-d que les éléments diagonaux de  $I$  sont égaux à 1 et les autres sont égaux à 0.

$$I = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \vdots & \vdots & 1 & 0 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix} \quad (5.10)$$

alors :

$$A.I = I.A = A \quad (5.11)$$

Donc  $I$  est l'élément neutre pour la multiplication des matrices.

## 5.7 PROPRIÉTÉS DES MATRICES

\* Matrice transposée : soit la matrice  $A = (a_{ij})$  et  $B = (b_{ij})$ .

$B$  est la matrice transposée de  $A \Leftrightarrow b_{ji} = a_{ij}$  pour  $i, j = 1, 2, \dots, n$  et on note  $B^T$ .

\* Matrice symétrique, soit  $A = (a_{ij})$ ,  $A$  est symétrique  $\Leftrightarrow A = A^T$  soit  $a_{ij} = a_{ji}$  pour  $i, j = 1, 2, \dots, n$ .

\* Matrice antisymétrique :  $A$  est antisymétrique  $A^T = -A$  soit  $a_{ii} = 0$  pour  $i = 1, 2, \dots, n$ .

\*  $A = (a_{ij})$  une matrice définie positive est une matrice carrée telle que :

$$\forall \text{ le vecteur } x \neq \mathbf{0} \Rightarrow x^T . A . x > 0. \quad (5.12)$$

\* Une matrice orthogonale est une matrice  $A$  où toutes les lignes ou colonnes  $c_i$  vérifient :

$c_i^T . c_j = 0$  si  $i \neq j$  et  $c_i^T . c_j = 1$  si  $i=j$ . Alors :

$$A^{-1} = A^T \quad (5.13)$$

\* Matrice diagonale  $A$  : elle s'écrit  $A = (a_{ii})_{i=1,2,\dots,n}$ .

Soit  $A$  la matrice :

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (5.14)$$

Alors le déterminant de  $A$  est égal à :

$$\text{dét}(A) = a.d - b.c \quad (5.15)$$

Et on note :

$$\text{Dét}(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \quad (5.16)$$

**Théorème 5.1** Si  $\text{dét}(A)$  est non nul, alors la matrice  $A$  est inversible.

Soit la matrice d'ordre 3 :

$$A = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} \quad (5.17)$$

Alors le déterminant de  $A$  est donné par :

$$\text{Dét}(A) = a \begin{vmatrix} b' & c' \\ b'' & c'' \end{vmatrix} - a' \begin{vmatrix} b & c \\ b'' & c'' \end{vmatrix} + a'' \begin{vmatrix} b & c \\ b' & c' \end{vmatrix} \quad (5.18)$$

On a alors :

$$\text{Dét}(A.B) = \text{Dét}(B.A) = \text{Dét}(A). \text{Dét}(B) \quad (5.19)$$

**Théorème 5.2** Une matrice carrée  $A$  est inversible s'il existe une matrice unique notée  $A^{-1}$  appelée matrice inverse de  $A$  telle que :  $A.A^{-1} = A^{-1}.A = I$

On a les propriétés suivantes :

$$(A.B)^{-1} = B^{-1}.A^{-1} \quad (5.20)$$

$$(A^T)^{-1} = (A^{-1})^T \quad (5.21)$$

**Définition 5.6** La trace d'une matrice  $A = (a_{ij})$  est donnée par :



$$\text{trace}(A) = \text{tr}(A) = \sum_{i=1}^{i=n} a_{ii} \quad (5.22)$$

D'où les propriétés :

$$\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B) \quad (5.23)$$

$$\text{tr}(A.B) = \text{tr}(B.A) \quad (5.24)$$

## 5.8 VALEURS PROPRES D'UNE MATRICE

**Définition 5.7** Les valeurs propres  $\lambda_i = \lambda_i(A), i = 1, 2, \dots, n$  d'une matrice  $A$  carrée d'ordre  $n$  sont les racines réelles ou complexes, distinctes ou confondues du polynôme caractéristique :

$$P_A : \lambda \in \mathbb{C} \longrightarrow P_A(\lambda) = \text{dét}(A - \lambda.I)$$

de la matrice  $A$ .

On a aussi les propriétés :

$$\text{tr}(A) = \sum_{i=1}^{i=n} \lambda_i \quad (5.25)$$

$$\text{Dét}(A) = \prod_{i=1}^{i=n} \lambda_i(A) \quad (5.26)$$

**Définition 5.8** On appelle rayon spectral d'une matrice  $A$  le nombre  $\rho \geq 0$  tel que :

$$\rho(A) = \max_{i \in [1, n]} \{|\lambda_i(A)|\} \quad (5.27)$$

### 5.8.1 Vecteurs propres

**Définition 5.9** A toute valeur propre  $\lambda_i$  d'une matrice carrée  $A$  est associée au moins un vecteur  $u \in \mathbb{R}^n$  tel que :

$$u \neq 0 \quad \text{et} \quad A.u = \lambda.u \quad (5.28)$$

## 5.9 RANG D'UNE APPLICATION LINÉAIRE

Soient  $U$  et  $V$  deux espaces vectoriels réels de dimension  $n$  et  $f : U \rightarrow V$  une application linéaire, d'où :

**Définition 5.10** *Le rang de l'application  $f : U \rightarrow V$  est égal à la dimension du sous-espace vectoriel image de  $f$  :*

$$\text{Im}(f) = \{f(u) \in V, \forall u \in U\}$$

Si  $f$  est représentée par une matrice  $A$ , le rang de  $f$  est égal au plus grand ordre des sous-matrices carrées inversibles de la matrice  $A$ . C'est pourquoi le rang de  $f$  c'est aussi le rang de  $A$  qu'on note  $\text{rang}(A)$  ou  $r(A)$ .

## 5.10 EXERCICES

### Exercice n°1 :

Démontrer que l'application  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  telle que  $(x, y, z) \rightarrow (y + z, x + z, x + y)$  est une application linéaire de  $\mathbb{R}^3 \rightarrow \mathbb{R}^3$ . Quel est son rang.

### Exercice n°2 :

$E$  désigne l'espace vectoriel engendré par les fonctions  $f_1(x) = \cos x$  et  $f_2(x) = \sin x$ . Soit  $T$  l'application linéaire de  $E$  dans lui même définie par  $T(f) = f'$  c'est-à-dire la dérivée de  $f$ . Trouver par rapport à la base  $(f_1, f_2)$  la matrice de  $T$ , puis de  $T^2 = T \circ T$ ,  $T^3$  et  $T^4$ .

### Exercice n°3 :

Déterminer le rang des matrices suivantes :

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 1 & 2 & 0 & 0 \end{pmatrix}.$$

**Exercice n°4 :**

On note  $M_n(\mathbb{R})$  l'ensemble des matrices carrées d'ordre  $n$  à coefficients réels et on rappelle qu'une norme matricielle sur  $M_n(\mathbb{R})$  est une application vérifiant les propriétés suivantes :

$$* \|A\| = 0 \iff A = 0, \text{ et } \|A\| \geq 0, \forall A \in M_n(\mathbb{R}),$$

$$* \|\alpha A\| = |\alpha| \|A\|, \forall \alpha \in \mathbb{R}, \forall A \in M_n(\mathbb{R}),$$

$$* \|A + B\| \leq \|A\| + \|B\|, \forall A, B \in M_n(\mathbb{R}).$$

On considère l'application définie par  $F : A = (a_{ij}) \in M_n(\mathbb{R}) \longrightarrow F(A) = \left( \sum_{i,j} a_{ij}^2 \right)^{1/2}$ .

1. Montrer que cette application est une norme matricielle sur  $M_n(\mathbb{R})$  qu'on note  $\|\cdot\|_F$ .

**Exercice n°5 :**

1. Calculer le déterminant de la matrice  $A$  :

$$A = \begin{pmatrix} 5 & -1 & 0 \\ 2 & 3 & 1 \\ -1 & 0 & 2 \end{pmatrix}$$

2. Ecrire le polynôme caractéristique  $P_A(\lambda)$ .

3. Déterminer les valeurs propres de  $A$ .

**Exercice n°6 :**

Pour tout réel  $\alpha$ , on note  $A_\alpha$  la matrice :

$$A_\alpha = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix}$$

1. Quelles sont les valeurs propres de la matrice  $A_\alpha$  ?

2. Pour quelles valeurs de  $\alpha$  la matrice  $A_\alpha$  est-elle inversible ?

**Exercice n°7 :**

Soit  $A = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix}$ .

1. Calculer  $A^T.A$ ,  $A$  est-elle inversible? Calculer son inverse.

**Exercice n°8 :**

Soit la matrice  $A$  :

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

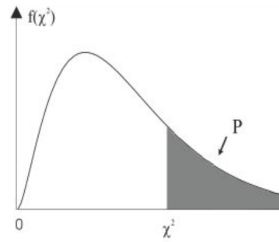
et  $J = A - I$  où  $I$  est la matrice unité d'ordre 3 :

1. Déterminer l'entier  $k$  tel que  $J^k = 0$ .
2. En déduire  $A^n$  pour tout entier  $\geq 3$ .

## Annexe B : Table du $\chi^2$

On présente dans la page suivante une table de la loi du  $\chi^2$  [3].

Valeurs de  $\chi^2$  ayant la probabilité P d'être dépassées



| $\nu$ | P = 0,995 | 0,99   | 0,975  | 0,95   | 0,90   | 0,10   | 0,05   | 0,025  | 0,01   | 0,005  |
|-------|-----------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1     | 0,0004    | 0,0002 | 0,001  | 0,0039 | 0,0158 | 2,706  | 3,841  | 5,024  | 6,635  | 7,879  |
| 2     | 0,010     | 0,020  | 0,051  | 0,103  | 0,211  | 4,605  | 5,991  | 7,378  | 9,210  | 10,597 |
| 3     | 0,072     | 0,115  | 0,216  | 0,352  | 0,584  | 6,251  | 7,815  | 9,348  | 11,345 | 12,838 |
| 4     | 0,207     | 0,297  | 0,484  | 0,711  | 1,064  | 7,779  | 9,488  | 11,143 | 13,277 | 14,860 |
| 5     | 0,412     | 0,554  | 0,831  | 1,145  | 1,610  | 9,236  | 11,070 | 12,833 | 15,086 | 16,750 |
| 6     | 0,676     | 0,872  | 1,237  | 1,635  | 2,204  | 10,645 | 12,592 | 14,449 | 16,812 | 18,548 |
| 7     | 0,989     | 1,239  | 1,690  | 2,167  | 2,833  | 12,017 | 14,067 | 16,013 | 18,475 | 20,278 |
| 8     | 1,344     | 1,646  | 2,180  | 2,733  | 3,490  | 13,362 | 15,507 | 17,535 | 20,090 | 21,965 |
| 9     | 1,735     | 2,088  | 2,700  | 3,325  | 4,168  | 14,684 | 16,919 | 19,023 | 21,666 | 23,589 |
| 10    | 2,156     | 2,558  | 3,247  | 3,940  | 4,865  | 15,987 | 18,307 | 20,483 | 23,209 | 25,188 |
| 11    | 2,603     | 3,053  | 3,816  | 4,575  | 5,578  | 17,275 | 19,675 | 21,920 | 24,725 | 26,757 |
| 12    | 3,074     | 3,571  | 4,404  | 5,226  | 6,304  | 18,549 | 21,026 | 23,337 | 26,217 | 28,300 |
| 13    | 3,565     | 4,107  | 5,009  | 5,892  | 7,042  | 19,812 | 22,362 | 24,736 | 27,688 | 29,819 |
| 14    | 4,075     | 4,660  | 5,629  | 6,571  | 7,790  | 21,064 | 23,685 | 26,119 | 29,141 | 31,319 |
| 15    | 4,601     | 5,229  | 6,262  | 7,261  | 8,547  | 22,307 | 24,996 | 27,488 | 30,578 | 32,801 |
| 16    | 5,142     | 5,812  | 6,908  | 7,962  | 9,312  | 23,542 | 26,296 | 28,845 | 32,000 | 34,267 |
| 17    | 5,697     | 6,408  | 7,564  | 8,672  | 10,085 | 24,769 | 27,587 | 30,191 | 33,409 | 35,718 |
| 18    | 6,265     | 7,015  | 8,231  | 9,39   | 10,865 | 25,989 | 28,869 | 31,526 | 34,805 | 37,156 |
| 19    | 6,844     | 7,633  | 8,907  | 10,117 | 11,651 | 27,204 | 30,144 | 32,852 | 36,191 | 38,582 |
| 20    | 7,434     | 8,260  | 9,591  | 10,851 | 12,443 | 28,412 | 31,410 | 34,170 | 37,566 | 39,997 |
| 21    | 8,034     | 8,897  | 10,283 | 11,591 | 13,240 | 29,615 | 32,671 | 35,479 | 38,932 | 41,401 |
| 22    | 8,643     | 9,542  | 10,982 | 12,338 | 14,041 | 30,813 | 33,924 | 36,781 | 40,289 | 42,796 |
| 23    | 9,260     | 10,196 | 11,689 | 13,091 | 14,848 | 32,007 | 35,172 | 38,076 | 41,638 | 44,181 |
| 24    | 9,886     | 10,856 | 12,401 | 13,848 | 15,659 | 33,196 | 36,415 | 39,364 | 42,980 | 45,559 |
| 25    | 10,520    | 11,524 | 13,120 | 14,611 | 16,473 | 34,382 | 37,652 | 40,646 | 44,314 | 46,928 |
| 26    | 11,160    | 12,198 | 13,844 | 15,379 | 17,292 | 35,563 | 38,885 | 41,923 | 45,642 | 48,290 |
| 27    | 11,808    | 12,879 | 14,573 | 16,151 | 18,114 | 36,741 | 40,113 | 43,195 | 46,963 | 49,645 |
| 28    | 12,461    | 13,565 | 15,308 | 16,928 | 18,939 | 37,916 | 41,337 | 44,461 | 48,278 | 50,993 |
| 29    | 13,121    | 14,256 | 16,047 | 17,708 | 19,768 | 39,087 | 42,557 | 45,722 | 49,588 | 52,336 |
| 30    | 13,787    | 14,953 | 16,791 | 18,493 | 20,599 | 40,256 | 43,773 | 46,979 | 50,892 | 53,672 |

Note : L'entier  $\nu$  est le nombre de degrés de liberté de  $\chi^2$ .

Fig. 5.1 Loi du Kui-deux  $\chi^2(\nu)$

# Bibliographie

1. **P. Hottier**. 1980. Théorie des erreurs. Ecole Nationale des Sciences Géographiques (ENSG). IGN France.
2. **P. Hottier**. 1980. Cours de statistique. ENSG. IGN France.
3. **C. Hardouin, A.K. Fermin**. Enquêtes et méthodes d'analyse quantitative. Master 1 SES. (consultation Internet, juillet 2020).
4. **Yu V. Linnik**. 1958. Méthode des Moindres Carrés et les Bases Mathématiques de la théorie Statistique et le traitement des Observations (en russe). Publications de Littérature, Physique et Mathématique de l'Etat. Moscou, 336 pages.





# Liste des figures

|     |   |     |
|-----|---|-----|
| 0.1 | Prof. Mohamed Rached Boussema (1953-2020).....                            | xii |
| 1.1 | L'Exemple Fondamental .....   | 5   |
| 1.2 | Le Choix des poids .....  | 10  |
| 2.1 | L'Exemple fondamental .....   | 18  |
| 2.2 | Etapes du calcul de la descente par l'algorithme de Gauss-Doolittle ..... | 49  |
| 2.3 | Etapes du calcul de la descente par l'algorithme de Cholewsky .....       | 52  |
| 2.4 | Méthode des équations de condition liant les grandeurs observées .....    | 58  |
| 2.5 | Equations d'observations + équations de condition.....                    | 60  |
| 3.1 | Allure du système normal .....  | 74  |
| 3.2 | Le problème du lac .....  | 76  |
| 4.1 | Ellipses de variance de $\tilde{X}$ et $\tilde{X}'$ .....                 | 96  |
| 4.2 | La déformée type 1 .....  | 99  |
| 4.3 | La déformée type 2 .....  | 99  |
| 5.1 | Loi du Kui-deux $\chi^2(v)$ .....   | 124 |