

Exemple d'Application de la Méthode des Moindres Carrés

Par
Abdelmajid BEN HADJ SALEM

INGÉNIEUR GÉNÉRAL GÉOGRAPHE RETRAITÉ DE L'OFFICE DE LA
TOPOGRAPHIE ET DU CADASTRE

JANVIER 2020

VERSION 1.

Email : abenhadsalem@gmail.com

Table des matières

| | | |
|----------|--|----------|
| 1 | Introduction | 3 |
| 1.1 | Notations | 3 |
| 2 | Définition du problème | 4 |
| 2.1 | Les Mesures | 4 |
| 3 | La Méthode des Moindres Carrés - Cas du Modèle Linéaire de Gauss-Markov | 5 |
| 3.1 | Solution par la Méthode des Moindres Carrés | 5 |
| 3.2 | La matrice de variance des inconnues | 7 |
| 4 | Application à un exemple | 7 |
| 4.1 | Notations et données | 7 |
| 4.2 | Calculs et Résultats | 7 |
| 5 | Références | 9 |

Exemple d'Application de la Méthode des Moindres Carrés

Abdelmajid Ben Hadj Salem

Résumé

Dans cette note, on présente la méthode des moindres carrés et son application dans les travaux topographiques

1 Introduction

Dans le métier de l'ingénieur géomètre, on rencontre des cas où on est amené à déterminer un ou plusieurs paramètres (altitude d'un point, les coordonnées d'un point, ...) en disposant d'un nombre d'observations supérieurs au nombre des inconnues. Dans ce cas, on applique la méthode des moindres carrés pour déterminer aux mieux nos inconnues.

1.1 Notations

Pour faciliter notre présentation de la méthode des moindres carrés, on considère qu'on veut déterminer 3 inconnues qu'on note X_1 , X_2 et X_3 . On présente ces vecteurs sous forme vectorielle par :

$$X = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \quad (1)$$

Les observations ou mesures sont représentés par le vecteur L de n composantes, avec n le nombres des observations. On donnera par la suite la définition des composantes l_i de L :

$$L = \begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix} \quad (2)$$

La norme (ou longueur) d'un vecteur Y est donnée par :

$$\|Y\| = \sqrt{Y \cdot Y} = \sqrt{Y^T \cdot Y} \implies \|Y\|^2 = Y^T \cdot Y = Y_1^2 + Y_2^2 + Y_3^2 \quad (3)$$

où T désigne transposée.

2 Définition du problème

On veut déterminer les 3 grandeurs scalaires inconnues X_1, X_2, X_3 à l'aide de n grandeurs observées distinctes ou non des précédentes, mais qui leur sont liées géométriquement (ou physiquement).

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 $3 - p$: et il n'y aura en fait que $3 - p$ inconnues scalaires. Une condition nécessaire pour qu'on puisse déterminer ces $3 - p$ inconnues est que $n \geq 3 - p$. Autrement dit il faut que les grandeurs observées déterminent géométriquement (ou physiquement) les grandeurs à déterminer.

2.1 Les Mesures

Les mesures l_1, l_2, \dots, l_n 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 suivent la loi normale.

La matrice variance du vecteur observation L est alors diagonale.

$$\Gamma_L = \begin{pmatrix} emq_1^2 & 0 & 0 & 0 \\ 0 & emq_2^2 & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & emq_n^2 \end{pmatrix} \quad (4)$$

- indirectes, dans ce cas, elles sont alors en principe généralement corrélées, corrélation qui se traduit par le fait que Γ_L n'est pas diagonale.

Dans tout ce qui suit, nous supposons en principe que les mesures sont normales. On a donc :

- les 3 grandeurs inconnues,
- les n grandeurs observées dans des conditions d'exactitude déterminée.

Le but que nous nous proposons est alors :

- d'estimer au mieux à partir de la donnée des n mesures la valeur des 3 grandeurs inconnues.
- de trouver le meilleur estimateur de la matrice variance du vecteur observation L , c'est-à-dire Γ_L .

- de trouver le meilleur estimateur de la matrice variance Γ_X de X qui chiffrera "la précision" de X .

3 La Méthode des Moindres Carrés - Cas du Modèle Linéaire de Gauss-Markov

Le modèle linéaire de Gauss-Markov est défini par :

$$\boxed{A.X = L - e; \quad e \in \mathcal{N}(0, \Gamma)} \quad (5)$$

avec :

- L : le vecteur des observations ($n \times 1$) = $(l_1, l_2, \dots, l_n)^T$,
- X : le vecteur des inconnues (3×1) = $(X_1, X_2, X_3)^T$,
- e : le vecteur des erreurs ($N \times 1$) = $(e_1, e_2, \dots, e_n)^T$ suit la loi normale $\mathcal{N}(0, \Gamma)$ avec $E(e) = 0$ (E désigne l'espérance mathématique ou moyenne) et $\Gamma = E(ee^T)$ la matrice de dispersion ou variance, on prendra $\Gamma = \sigma_0^2.P^{-1}$. P est la matrice des poids et σ_0 une constante positive. Au lieu de considérer le vecteur des erreurs e , on utilise le vecteur $V = -e$ appelé vecteur des résidus :

$$V = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \quad (6)$$

Alors l'équation (5) devient :

$$\boxed{A.X = L + V} \quad (7)$$

Dans notre note on prendra $\sigma_0^2 = 1$ et P la matrice unité 3×3 :

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (8)$$

- A : est la matrice des constantes, c'est une matrice $n \times 3$.

3.1 Solution par la Méthode des Moindres Carrés

Alors la solution par les moindres carrés \bar{X} sera définie par :

$$\|A\bar{X} - L\|^2 = \min \|A.X - L\|^2 = \min \|V\|^2 = \sum_{i=1, n} v_i^2 \quad (9)$$

En effet, on veut minimiser la fonction :

$$F(X) = F(X_1, X_2, X_3) = \|A.X - L\|^2 = \|A.X(X_1, X_2, X_3) - L\|^2 \quad (10)$$

On obtient :

$$\frac{\partial F(X)}{\partial X} = 0$$

soit :

$$\frac{\partial}{\partial X} \|A.X - L\|^2 = 0 \quad (11)$$

or :

$$\|A.X - L\|^2 = (A.X - L)^T . (A.X - L) = X^T . A^T A.X - 2L^T . AX + L^T . L \quad (12)$$

Posons :

$$\boxed{N = A^T . A} \quad (13)$$

N est appelée matrice normale. Soit :

$$\begin{aligned} \frac{\partial F(X)}{\partial X} &= 2X^T . (A^T . A)X - 2X^T . A^T . L = 2X^T . N.X - 2X^T . A^T . L \implies \\ \frac{\partial F(X)}{\partial X} &= 0 \implies X^T . N.X - X^T . A^T . L = 0 \implies \\ &X^T . N.X = A^T . L \end{aligned} \quad (14)$$

Alors la matrice N est symétrique ($N^T = N$) et inversible c'est-à-dire régulière et de rang $r = 3$. En effet la matrice $A =_n A_r$ est de rang $r = 3$ c'est-à-dire on peut extraire une sous matrice ${}_r A'_r$ de A telle que son déterminant est différent de zéro \implies Déterminant(A') $\neq 0$, de plus la matrice N est définie positive, on entend par là que $X \neq 0 \implies X^T . N.X > 0$ car :

$$X^T . N.X = (X^T A^T) . (AX) = (AX)^T . (AX) = \|AX\|^2 > 0 \quad (15)$$

le carré de la norme du vecteur $A.X$ et pour $X \neq 0$, on a $A.X \neq 0$ sinon la matrice A serait de *rang* < 3 , par suite (15) est vérifiée. Le vecteur \bar{X} solution des moindres carrés est donné par :

$$\boxed{\bar{X} = (A^T A)^{-1} A^T L = N^{-1} A^T L} \quad (16)$$

Le vecteur des observations compensé est donné par :

$$\boxed{\bar{L} = A\bar{X}} \quad (17)$$

Le vecteur des résidus est obtenu par :

$$\boxed{V = A\bar{X} - L = AN^{-1}A^T L - L = (AN^{-1}A^T - I).L} \quad (18)$$

Pour vérifier la solution des moindres carrés, on calcule :

$$\begin{aligned} A^T V &= A^T (AN^{-1}A^T - I).L = A^T AN^{-1}A^T L - A^T L = A^T L - A^T L = 0 \\ &\boxed{A^T . L = 0} \end{aligned} \quad (19)$$

soit $A^T L \equiv 0 =$ une matrice nulle. La condition (19) est appelée renormalisation. Elle est importante car elle garantit que le résultat obtenu \bar{X} est bien celui des moindres carrés.

3.2 La matrice de variance des inconnues

Une estimation de la matrice variance de \bar{X} est donnée par :

$$\Gamma_{\bar{X}} = s_0^2 N^{-1}, \quad s_0^2 = \frac{\|V\|^2}{n-r} = \frac{\sum_{i=1,n} v_i^2}{n-3} \quad (20)$$

4 Application à un exemple

On dispose des coordonnées planimétriques $(X_i, Y_i), i = 1, n$ de n points situés uniformément distribués le long du périmètre d'une citerne quasicirculaire. Le problème est de déterminer le cercle (coordonnées du centre et le rayon) qui épouse au mieux la citerne.

4.1 Notations et données

Soient (X_0, Y_0) les coordonnées approchées du centre qu'on peut calculer comme :

$$X_0 = \frac{\sum_{i=1}^n X_i}{n} \quad (21)$$

$$Y_0 = \frac{\sum_{i=1}^n Y_i}{n} \quad (22)$$

Les inconnues pour le point centre sont dX, dY , les coordonnées définitives seront :

$$X_c = X_0 + dX, \quad Y_c = Y_0 + dY$$

Le rayon approché R_0 peut être calculé par :

$$R_0 = \frac{\sum_{i=1,n} R_i}{n} = \frac{\sum_{i=1}^n \sqrt{(X_0 - X_i)^2 + (Y_0 - Y_i)^2}}{n}$$

Soit dR la correction à ajouter à la valeur approchée du rayon. Le rayon définitif $R = R_0 + dR$.

4.2 Calculs et Résultats

Les inconnues X, Y, R sont liées par l'équation :

$$(X - X_i)^2 + (Y - Y_i)^2 = R^2 \implies F(X, Y, R) = (X - X_i)^2 + (Y - Y_i)^2 - R^2 = 0, i = 1, n \quad (23)$$

On peut écrire l'équation précédente au deuxième ordre près comme suit :

$$\begin{aligned} F(X_0 + dX, Y_0 + dY, R_0 + dR) = 0 = & F(X_0, Y_0, R_0) + \frac{\partial F}{\partial X}(X_0, Y_0, R_0)dX + \\ & \frac{\partial F}{\partial Y}(X_0, Y_0, R_0)dY + \frac{\partial F}{\partial R}(X_0, Y_0, R_0)dR + \\ & \sum_{T=X,Y,R} \frac{1}{2} \frac{\partial^2 F}{\partial T^2}(X_0, Y_0, R_0)dT^2 \end{aligned} \quad (24)$$

Pour alléger la notation des équations, on note $C_0 = (X_0, Y_0, R_0)$. On peut écrire au premier ordre près :

$$\frac{\partial F_i}{\partial X}(C_0)dX + \frac{\partial F_i}{\partial Y}(C_0)dY + \frac{\partial F_i}{\partial R}(C_0) = -F_i(C_0) + v_i \quad (25)$$

C'est l'équation d'observations par point i , $i = 1, n$. Détaillons les coefficients des inconnues et le terme constant. D'où :

$$a_{i1} = \frac{\partial F_i}{\partial X}(C_0) = 2(X_0 - X_i) \quad (26)$$

$$a_{i2} = \frac{\partial F_i}{\partial Y}(C_0) = 2(Y_0 - Y_i) \quad (27)$$

$$a_{i3} = \frac{\partial F_i}{\partial R}(C_0) = -2R_i \quad (28)$$

$$l_i = -F_i(C_0) = -(R_i^2 - R_0^2) = R_0^2 - R_i^2 \quad (29)$$

On obtient alors l'équation (7) $A.X = L + V$:

$$\begin{pmatrix} 2(X_0 - X_1) & 2(Y_0 - Y_1) & -2R_1 \\ 2(X_0 - X_2) & 2(Y_0 - Y_2) & -2R_2 \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ 2(X_0 - X_n) & 2(Y_0 - Y_n) & -2R_n \end{pmatrix} \cdot \begin{pmatrix} dX \\ dY \\ dR \end{pmatrix} = \begin{pmatrix} R_0^2 - R_1^2 \\ R_0^2 - R_2^2 \\ \vdots \\ \vdots \\ R_0^2 - R_n^2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ \vdots \\ v_n \end{pmatrix} \quad (30)$$

Le vecteur $\bar{X} = (dX, dY, dR)^T$ est obtenu par :

$$\bar{X} = N^{-1}.A^T.L$$

On obtient les résultats définitifs :

$$X_c = X_0 + dX \quad (31)$$

$$Y_c = Y_0 + dY \quad (32)$$

$$R = R_0 + dR \quad (33)$$

Le calcul de la matrice normale $N = A^T.A$ donne :

$$N = \begin{pmatrix} \sum_{i=1,n} a_{i1}^2 & \sum_{i=1,n} a_{i1}a_{i2} & \sum_{i=1,n} a_{i1}a_{i3} \\ \sum_{i=1,n} a_{i1}a_{i2} & \sum_{i=1,n} a_{i2}^2 & \sum_{i=1,n} a_{i2}a_{i3} \\ \sum_{i=1,n} a_{i1}a_{i3} & \sum_{i=1,n} a_{i2}a_{i3} & \sum_{i=1,n} a_{i3}^2 \end{pmatrix} \quad (34)$$

Le vecteur $A^T.L$ est donné par :

$$A^T.L = \begin{pmatrix} \sum_{i=1,n} a_{i1}.l_i \\ \sum_{i=1,n} a_{i2}.l_i \\ \sum_{i=1,n} a_{i3}.l_i \end{pmatrix} \quad (35)$$

La condition de renormalisation sera vérifiée si $A^T V = 0$:

$$A^T \cdot V = \begin{pmatrix} \sum_{i=1,n} a_{i1} \cdot v_i \\ \sum_{i=1,n} a_{i2} \cdot v_i \\ \sum_{i=1,n} a_{i3} \cdot v_i \end{pmatrix} = (0, 0, 0)^T \quad (36)$$

Si $A^T L$ n'est pas nulle, dans ce cas on prendra comme vecteur approché $X_0 = \bar{X} + X_0 = (X_c, Y_c, R)^T$ et on réitérera le processus.

5 Références

Abdelmajid Ben Hadj Salem. *Eléments de Géodésie et de la Théorie des Moindres Carrés, pour les Elèves-Ingénieurs Géomaticiens.* ISBN : 978-3-330-96843-1. 364 pages. 2017. Noor Publishing.